Acta Psychologica

EUROPEAN JOURNAL OF PSYCHOLOGY

Le

EDITORS

H. C. J. DUIJKER (AMSTERDAM)
M. J. LANGEVELD (UTRECHT)
C. F. VAN PARREREN (AMSTERDAM)
J. Th. SNIJDERS (GRONINGEN)

AND

O. KLINEBERG (NEW YORK)
SECRETARY-GENERAL OF THE
INTERNATIONAL UNION OF SCIENTIFIC PSYCHOLOGY

PURDUE UNIVERSITY
JAN 29 1957
LIBRARY

Volume XII — No. 4 (1956)

SCANDINAVIAN ISSUE

NORTH-HOLLAND PUBLISHING COMPANY - AMSTERDAM

NATIONAL EDITORS

AUSTRIA	Dr. J. KOHLER *Innsbruck*
BELGIUM	Prof. J. NUTTIN *Leuven*
DENMARK	H. CHR. RASMUSSEN *København*
FINLAND	Dr. J. VON WRIGHT *Helsinki*
FRANCE	Prof. P. FRAISSE *Paris*
GERMANY	Prof. A. WELLEK *Mainz*
GREAT BRITAIN	Prof. J. COHEN *Manchester*
ITALY	Prof. C. L. MUSATTI *Milano*
NORWAY	Dr. E. ØSTLYNGEN *Oslo*
SWEDEN	Dr. C. I. SANDSTRÖM *Stockholm*
SWITZERLAND	Prof. B. INHELDER *Genève*

CONTENTS:

Gösta Ekman: Discriminal sensitivity on the subjective continuum 233

Kullervo Rainio: The effect of the selection situation on responses to questionnaires 244

C. I. Sandström and *I. Lundberg:* A genetic approach to sex differences in localization 247

Per Saugstad: Concepts in psychological model construction 254

Gudmund Smith and *Maj Henriksson:* Studies in the development of a percept within various contexts of perceived reality 263

Alexander W. Stern: The nature of G and the concept of intelligence 282

J. M. von Wright: On the 'correction' and 'non-correction' methods of learning 290

This issue is published in cooperation with *Nordisk Psykologi.*

DISCRIMINAL SENSITIVITY ON THE SUBJECTIVE CONTINUUM

GÖSTA EKMAN
University of Stockholm, Sweden

INTRODUCTION

In a classic paper from 1927, Thurstone laid down some fundamental principles of modern psychophysics (13). He derived the law of comparative judgment (14), which refers to the subjective continuum and is expressed in subjective units (13). Thurstone applied his concepts to the Weber and Fechner laws, demonstrating that these laws would be valid together only on the condition that the discriminal dispersion is constant over the subjective continuum (13).

The law of comparative judgment and related concepts have usually been applied to scaling problems where no simple physical dimension is involved. In scaling methods which are based on this law, the relation of the discriminal dispersion to scale values on the subjective continuum is an important question for methodological reasons, but there is no way of investigating this relation directly, since the scale values are derived from discriminal dispersions.

The discriminal sensitivity on the subjective continuum and its relation to subjective scale values seems, however, to be an intrinsically interesting problem. It is the purpose of this paper to draw attention to this problem and to demonstrate how the relation may be investigated in psychophysical studies where a well-defined physical continuum is involved.

PROBLEM AND METHOD

A stimulus S_j gives rise to a distribution of discriminal processes R_{hj} with the modal process R_j, where R denotes perceived intensity on a homogeneous psychological continuum. These concepts are illustrated in Fig. 1.

The form of the frequency distribution connected with R_j may be investigated by determining the ordinates

$$y_{Rhj} = f_y(R_j). \qquad (1)$$

We make no assumptions concerning the form of this distribution. It may be studied by obtaining estimates of the ordinates according to a procedure to be described below.

This paper is an outcome of several discussions in a seminar on scaling theory. Special acknowledgements are due to contributions by Mr. M. Björkman and Dr. G. Johansson.

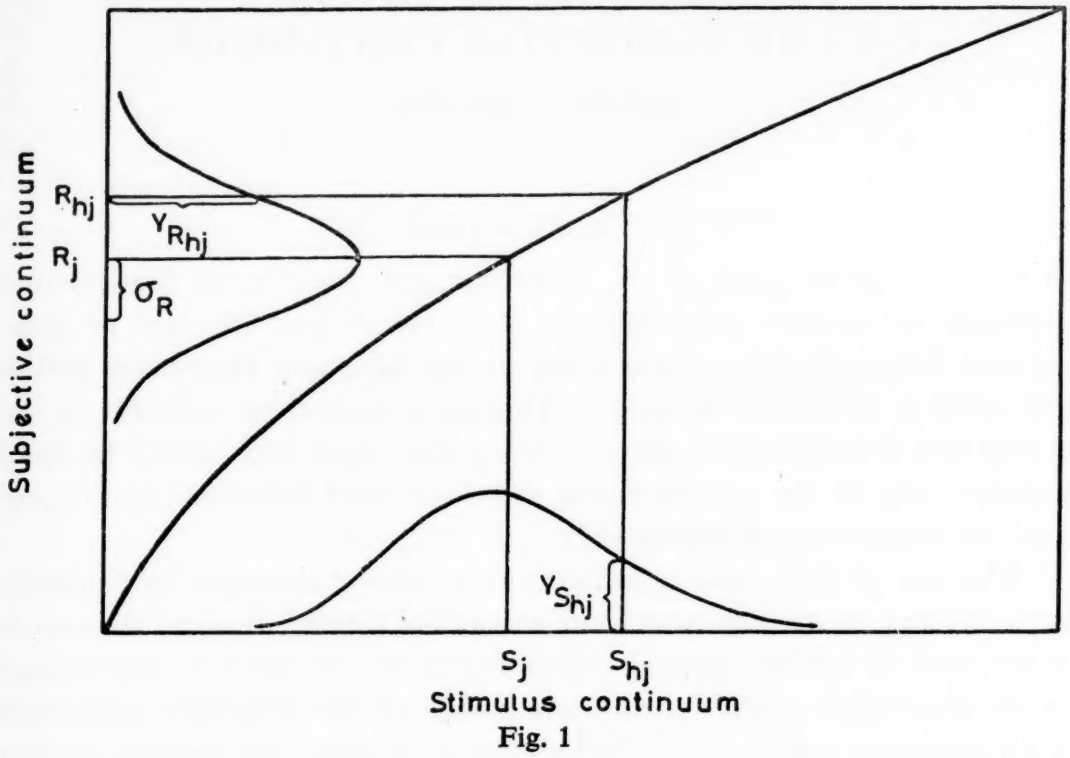


Fig. 1

In general, however, we are not primarily interested in the form of the distribution, but in obtaining a measure of its dispersion, which is an inverse measure of the discriminal sensitivity on the subjective continuum. Our problem, then, is to investigate the relation between discriminal dispersion and perceived intensity. In general, this relation is expressed by the equation

$$\sigma_R = f_\sigma (R), \tag{2}$$

where the particular function and its constants are to be experimentally determined. This equation is entirely in terms of R and contains no reference to S.

As an intermediate step, the physical continuum S has to be introduced. The relation between R and S may, quite generally, be written

$$R = f_R(S). \tag{3}$$

The particular function relating R to S is not assumed to be mathematically defined. The values of R, which correspond to the experimentally manipulated variable S, may be determined by any convenient method not based on the law of comparative judgment. Conventional methods of equal-appearing intervals, or methods of fractionation or multiple stimuli, may be applied for obtaining the necessary scale values. The outcome of the scaling procedure will be a curve more or less similar to the one shown in Fig. 1.

In our illustration, the frequency distribution of discriminal processes is normal, and the corresponding distribution in experimentally observable units of S is slightly positively skewed. The two frequency functions, with ordinates y_R and y_S respectively, are related by the equation

$$y_R = \frac{dS}{dR} y_S. \tag{4}$$

When an estimate of dS/dR may be obtained (see below), it is possible to infer the form of the distribution of discriminal processes. From this distribution, σ_R may be computed.

In general, however, the range of discriminal processes and corresponding observations on S is so narrow that the curve relating R to S in this range comes very close to a straight line. The linear approximation will be very good in the case of difference limens in all sense modalities. Thus in such a small range S may be treated as a linear transformation of R, and the discriminal dispersion may be obtained from the equation

$$\sigma_R = \frac{dR}{dS} \sigma_S. \tag{5}$$

In order to find numerical values of σ_R, we have to find the derivative dR/dS. This is of course accomplished by differentiating equation (3), if the particular function is known, but an estimate of the derivative may also be obtained in other cases. A graphic procedure will be illustrated in the following section.

Once the numerical values of σ_R have been computed, they are plotted against R and the curve resulting from this plot is examined.

A NUMERICAL ILLUSTRATION

We may assume that we have obtained the curve shown in Fig. 2 as a result of a scaling procedure. The mathematical function relating R to S is not assumed to be known.

The process of graphic differentiation of R with respect to S starts with marking off the increments on R which correspond to a certain small increment of S. In our illustration, readings on R have been registered for intervals of five S units. The readings have been made from a graph 15×10 cm and may be considered relatively rough. The successive differences of these values have been divided by five, which is the number of S units in each interval. These quotients are rough estimates of dR/dS in the midpoint of the respective interval. They have been plotted in Fig. 3.

A smooth curve is drawn through the plot. This curve represents a better estimate of dR/dS. Readings from this curve are entered in the table

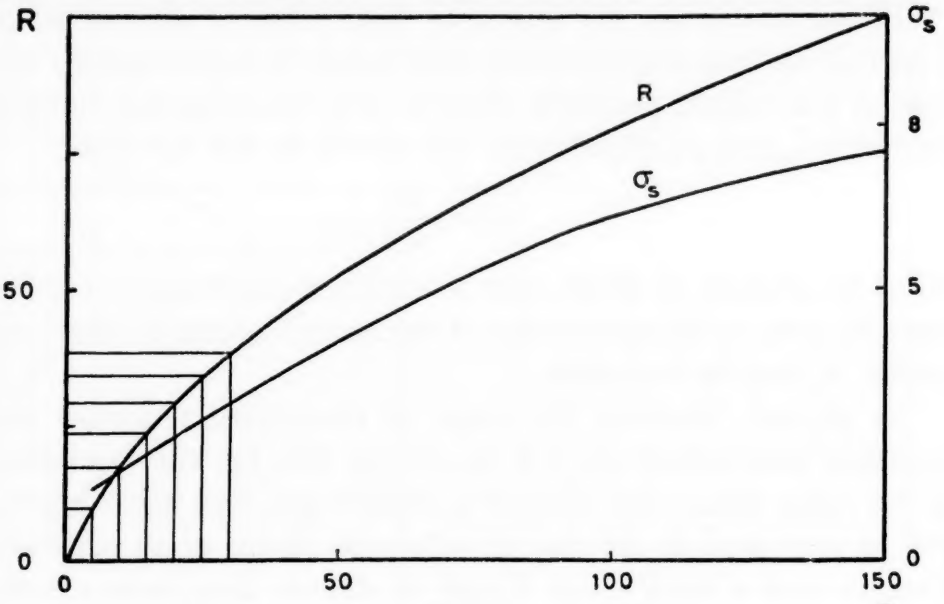


Fig. 2

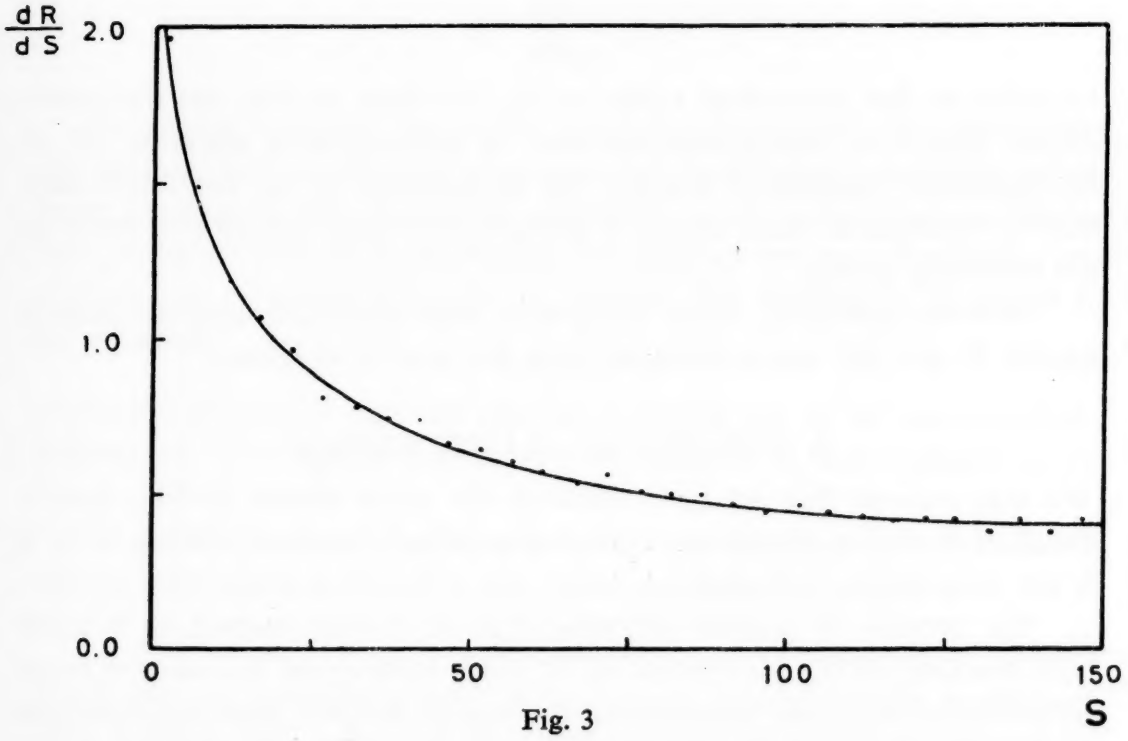


Fig. 3

below. This table is a numerical illustration of the application of equation (5) for computation of discriminal dispersions.

Finally, the values of σ_R are plotted against R, as in Fig. 4. In our example, the plot is linear, and a straight line with the equation $2 + 0.01$ R may be fitted to the plot. This, in fact, is the equation from which the values of σ_S in table 1 were constructed.

TABLE 1

S	R	σ_S	dR/dS	σ_R
5	9.8	1.31	1.60	2.10
10	17.2	1.67	1.30	2.17
25	33.8	2.63	0.89	2.34
50	52.6	4.01	0.63	2.53
75	67.1	5.29	0.50	2.65
100	79.1	6.31	0.44	2.78
125	89.9	7.04	0.41	2.89
150	100.0	7.54	0.40	3.02

The small deviations from the straight line in Fig. 4 are due to errors in the procedure of graphic differentation. These errors may be kept within considerably narrower limits when readings are made from graphs drawn on a larger scale.

APPLICATIONS TO EXPERIMENTAL DATA

In general, scaling and the determination of difference limens have been considered separate objects of investigation. The argument of this paper implies that difference limens or related measures are to be converted into R units and studied as a function of R. This, of course, makes it important

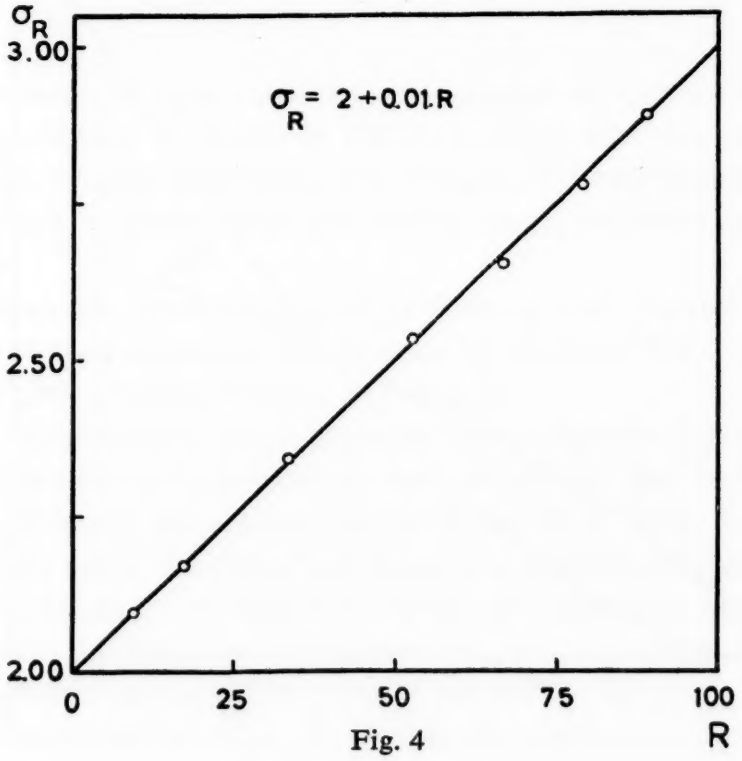


Fig. 4

to obtain both kinds of measures from the same subjects and, as far as possible, with the same experimental technique. Some work meeting with these requirements is now in progress at our laboratory.

By bringing together experimental data from separate investigations it may be possible to throw some light on our problem, though it must be remembered that because of technical and individual differences these data are not strictly comparable.

Weight

A scale of subjective weight was established by Harper and Stevens (7). The curve relating vegs to grams is shown in Fig. 5, together with the derivative obtained by the graphic procedure described above.

The authors converted DL's from Oberlin (10) into vegs, apparently by means of direct readings on R, and compared them with the S scale. We have converted Oberlin's data by equation (5) and plotted them against R in Fig. 6. The data from Oberlin's two experiments have been kept separate, since they were obtained from different subjects and under different conditions. When treated in this way, both sets of plotted values exhibit a general linear trend.

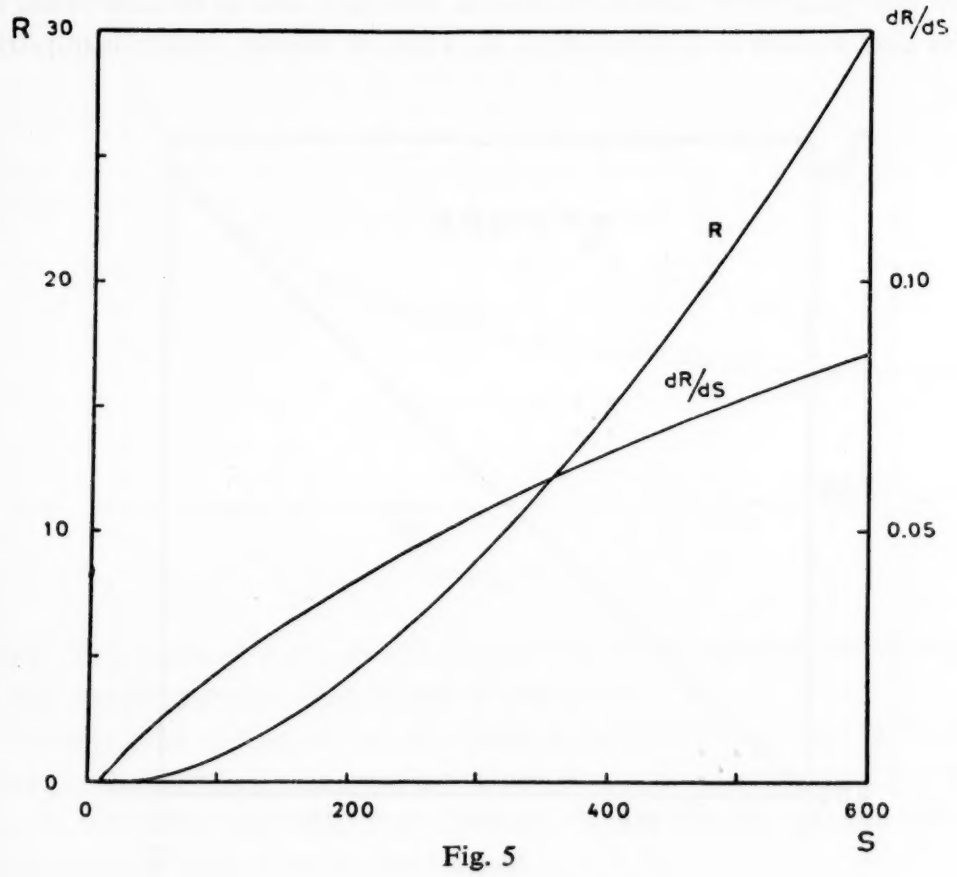


Fig. 5

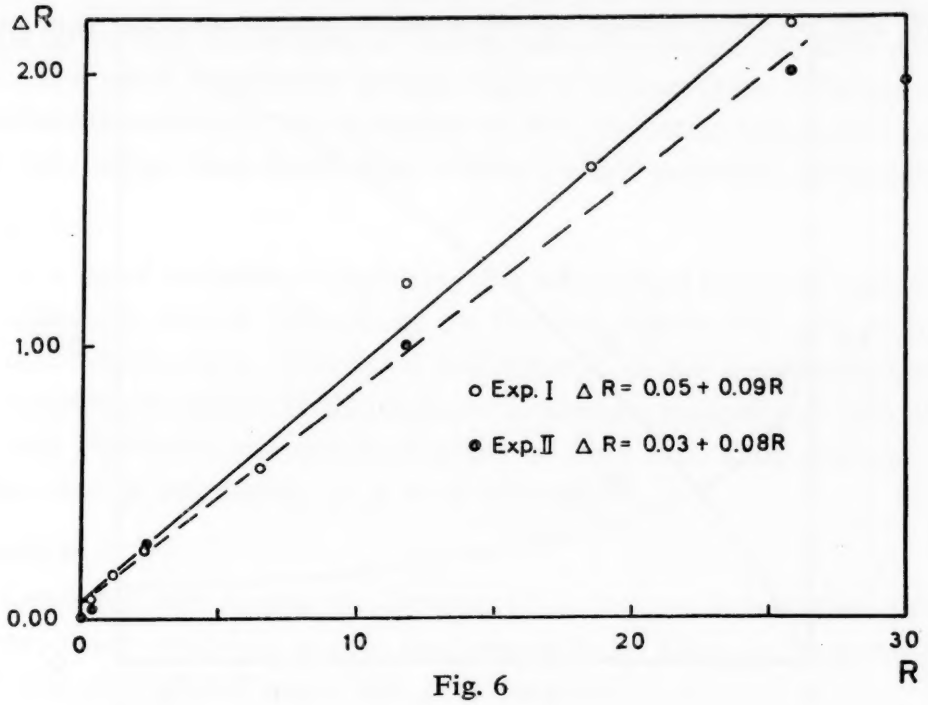


Fig. 6

In weight-lifting, the discriminal sensitivity in subjective units seems to decrease with subjective intensity. But the relative sensitivity is approximately constant. In other words, Weber's law applies as a fairly good approximation to the *subjective* dimension, but no far-reaching conclusions can be drawn, since the data brought together are not strictly comparable.

Taste

The intensity of taste sensations was scaled by Lewis (9). His curve for sodium chloride is shown in Fig. 7, together with the derivative obtained by the graphic procedure. The S scale of concentration, which was given by Lewis in grams solute per 100 cc water, has been converted into molar units.

The difference limens obtained by Holway and Hurvich (8) for two subjects have been converted into R units by means of Fig. 7 and equation (5) and are plotted against Lewis' R in Fig. 8.

Again, the difference limen increases with subjective intensity, i. e. the differential sensitivity in subjective units decreases. But in this case the relation is definitely non-linear. Up to about 70 R units, corresponding to three molar units, the difference limen is a retarded increasing function of subjective intensity. At higher R values, the difference limen again increases rapidly. It should be noted, however, that there were many reports of "sting" around the inflexion point of the curve. This may indicate that with increasing concentration, the subjective continuum of salt intensity

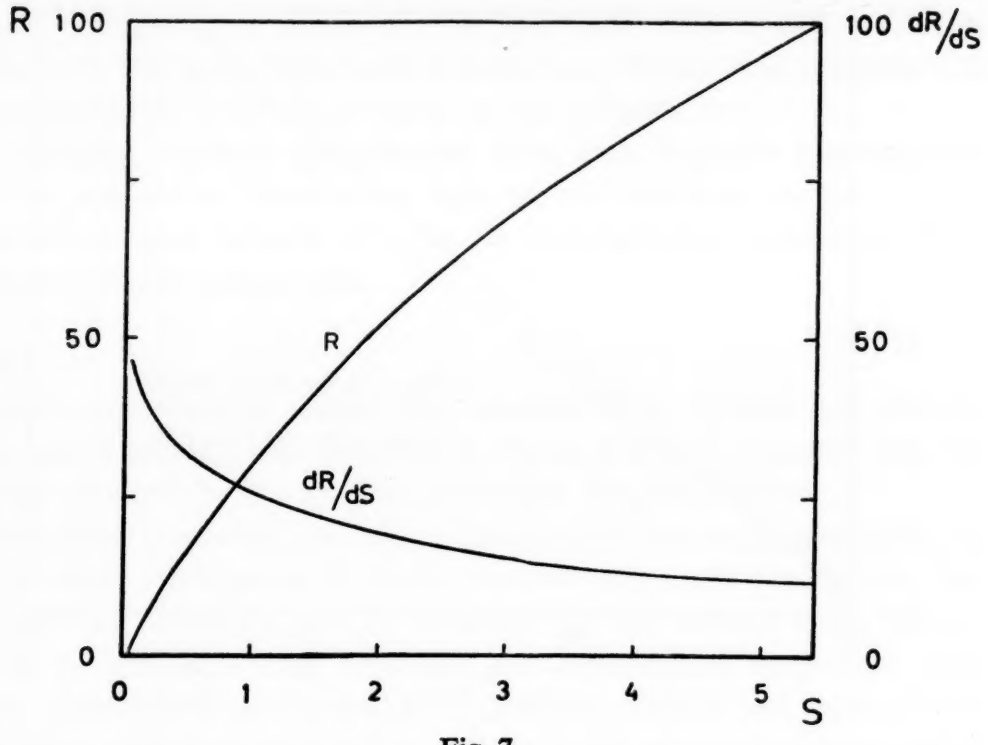

R
100
50
0
100 dR/dS
50
0
R
dR/dS
0
1
2
3
4
5
S

Fig. 7

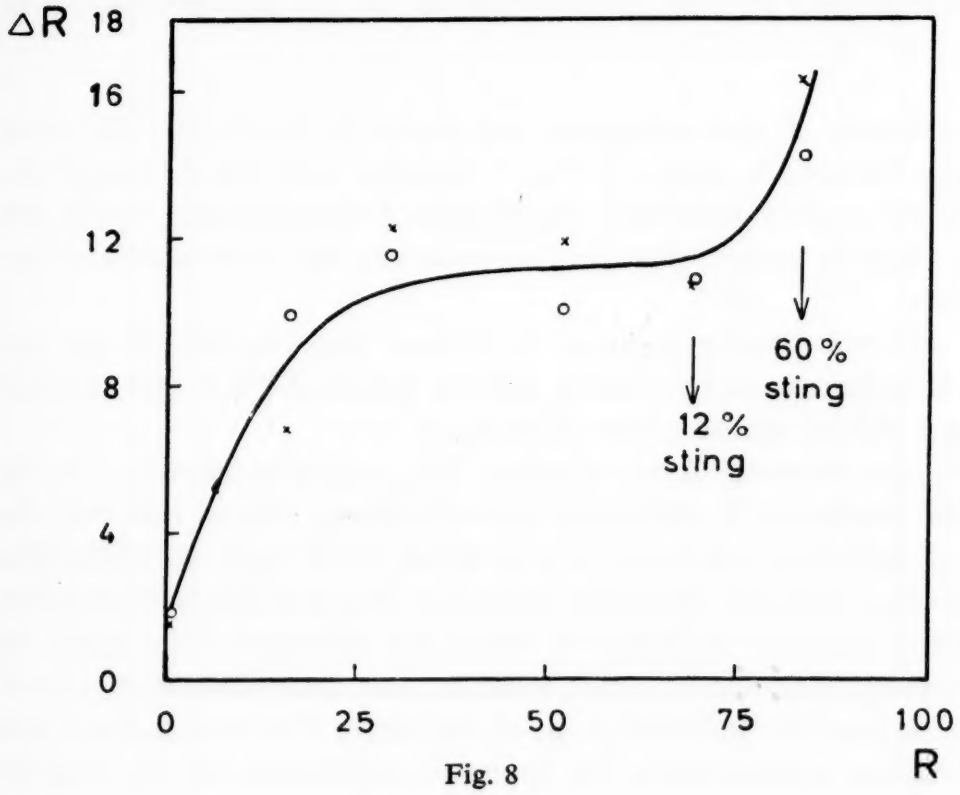

ΔR
18
16
12
8
4
0
0
25
50
75
100
R
60 %
sting
12 %
sting

Fig. 8

turns into a new dimension, involving pain, discomfort or other new qualities. The main conclusion is that there is a non-linear relation between discriminal sensitivity and intensity on the subjective continuum, as far as one may judge from these data, which are not perfectly comparable.

Brightness

A scale of subjective brightness was established by Hanes (6). Utilizing the difference limens determined by Troland, Hanes was able to integrate his curve from DL's. This is an indirect test of the hypothesis that DL's are subjectively equal. In this case, as Thurstone pointed out (13), Weber's law and Fechner's law will both apply or both fail. They are not verified in the case of brigthness, as is well known.

Loudness

Loudness was scaled by Stevens (11). It was found that difference limens, when scaled in sones, increase with S. Thus the hypothesis that DL's are subjectively equal has to be rejected.

Pitch

Stevens and Volkmann established a revised scale of pitch (12) and tried to predict DL's in terms of frequency, assuming them to be equal in terms of mels. Though there were some discrepancies between theoretical and experimental data, they concluded "that all jnd's for pitch are essentially equal in subjective size".

THE CASE OF QUALITATIVE VARIATION

So far, we have been concerned only with perceived position in a homogeneous subjective dimension. The instances of qualitative or inter-dimensional variation, however, are of considerable psychological interest and attract increasing attention.

When it is possible to trace the psychological correlate of a physical stimulus in a multidimensional subjective space, it is also possible to determine discriminal sensitivity as a function of the constants locating the experience in the subjective frame of reference.

In the field of color vision there is an indirect test of the hypothesis that discriminal sensitivity in a qualitative variation is constant along the curvilinear R continuum in the subjective space.

Hue

Hue was investigated in our laboratory by means of a method of similarity analysis (1). Five color factors were found to account for differ-

ences in hue. In a subsequent investigation it was shown that these color factors could be adequately described by normal curves on the base line of wave length (2), and finally we tried to predict DL's in terms of wave length on the assumption that discriminal sensitivity is constant in terms of units of the five subjective dimensions which are related by these normal curves to the physical continuum (3). The prediction was in good agreement with experimental data. As in all other cases, however, DL's are not strictly comparable with scaling data.

DISCUSSION

No obvious conclusions can be drawn from our survey of experiments. We have found three different types of relations between difference limen and position on the subjective continuum, namely (1) indications of a constant limen in brigthness and pitch, and in the qualitative variable of hue, (2) a linear relation in subjective weight, and (3) a definitely non-linear relation in salt intensity. All these comparisons are imperfect, however, since they have been made possible by bringing together data from different investigations. It is quite clear that difference limens and scale values ought to be determined by the same investigation.

Stevens and his coworkers have advanced the hypothesis that just noticeable differences are subjectively equal when variation in R depends on a change of structure excited, as in pitch, whereas they may increase when stimulation is added to stimulation already existing (7,12). Our own investigations in color vision lend some support to this hypothesis, but the constant sensitivity in brightness seems to contradict it.

Attempts to relate discriminal sensitivity or subjective intensity to the physical continuum, as in the Weber and Fechner laws or alternative psychophysical laws (4,5), are not too promising. There are too many peripheral physiological factors involved to make it probable that relatively simple and somewhat general relations can be found. Attempts to relate subjective discriminal sensitivity to subjective intensity, on the other hand, appear somewhat more promising. The peripheral factors are more or less ruled out by being present in both variables. Any relation may in this case possibly be ascribed to *central processes,* whereas the former relations involve both central *and* peripheral factors. This seems to be a rather good reason why we should investigate the relations between the sensitivity and position in the subjective continuum, including the multidimensional case.

SUMMARY

(1) It is pointed out that the relation between discriminal sensitivity and perceived intensity in the subjective dimension is an important problem which has attracted too little attention. There seems to be no single investigation directly centering on this question.

(2) A method is developed for investigating this relation. It is essentially a graphic procedure for differentiating R with respect to S in order to make possible a computation of discriminal dispersions or related measures in terms of R.

(3) The method is applied to a fictitious numerical example.

(4) The same method is applied to published experimental data. It is found that the difference limen is an increasing linear function of intensity in the case of subjective weight, and an increasing, definitely non-linear function in the case of subjective salt intensity.

(5) There is indirect evidence that the difference limen in subjective units is constant in the case of brightness and pitch, but is an increasing function of intensity in the case of loudness.

(6) The case of qualitative or inter-dimensional variation is discussed. There is indirect evidence that subjective sensitivity is constant along the curvilinear continuum in the multidimensional subjective space of hue.

(7) An hypothesis advanced by Stevens and collaborators is contradicted by brightness data but supported by color data.

(8) Relations between sensitivity and position in the subjective continuum may possible be ascribed to central processes, which makes them more promising objects of investigation than conventional psychophysical relations which also depend on highly specific peripheral factors.

REFERENCES

1. Ekman, G., Dimensions of color vision. *J. Psychol.*, 1954, *38*, 467-474.
2. – A note on the normal probability function in color vision. *J. Psychol.*, 1956, *41*, 231-234.
3. – Discrimination of hue as a function of wave length. *Acta psychol.*, 1956, *12*, 15-18 & *Nord. Psykol.*, 1956, *8*, 15-18.
4. Guilford, J. P., A generalized psychophysical law. *Psychol. Rev.*, 1932, *39*, 73-85.
5. – *Psychometric methods* (2nd ed.). N. Y.: McGraw-Hill, 1954.
6. Hanes, R. M., A scale of subjective brightness. *J. exp. Psychol.*, 1949, *39*, 438-452.
7. Harper, R. S., & Stevens, S. S., A psychological scale of weight and a formula for its derivation. *Amer. J. Psychol.*, 1948, *61*, 343-351.
8. Holway, A. H., & Hurvich, L. M., Differential gustatory sensitivity to salt. *Amer. J. Psychol.*, 1937, *49*, 37-48.
9. Lewis, D. R., Psychological scales of taste. *J. Psychol.*, 1948, *26*, 437-446.
10. Oberlin, K. W., Variations in intensitive sensitivity to lifted weights. *J. Exp. Psychol.*, 1936, *19*, 438-455.
11. Stevens, S. S., A scale for the measurement of a psychological magnitude: loudness. *Psychol. Rev.*, 1936, *43*, 405-416.
12. Stevens, S. S., & Volkmann, J., The relation of pitch to frequency: a revised scale. *Amer. J. Psychol.*, 1940, *53*, 329-353.
13. Thurstone, L. L., Psychophysical analysis. *Amer. J. Psychol.*, 1927, *38*, 368-389.
14. – A law of comparative judgment. *Psychol. Rev.*, 1927, *34*, 273-286.

THE EFFECT OF THE SELECTION SITUATION ON RESPONSES TO QUESTIONNAIRES

KULLERVO RAINIO
Institute of Industrial Supervision, Helsinki

One frequently used method of building a test battery for personnel selection purposes is to compare specified test variables with an acceptable criterion, such as performance on the job. In many cases the simplest way of evaluating the validity of a test is to examine workers already on the job for whom merit ratings or the like can be used as criteria. When such tests are used to select *new* applicants for similar jobs a new question arises: To what extent are the results affected by the applicant's desire to succeed? This question seems particularly important when questionnaires involving simple choices are used, where the results may depend not so much on individual ability as on personal motives. Thus there is reason to suppose that when applying for a job responses to simple choices may be consciously or unconsciously influenced by a desire to succeed in the selection situation, whereas for workers already employed a more "neutral" attitude may be assumed to provide greater variability of motives.

METHOD

In order to investigate this problem the performance of foremen already employed was compared with that of applicants for such jobs on 7 of the personality factors in the Thurstone Temperament Schedule (TTS) (7) and on the neuroticism variable of the Maudsley Medical Questionnaire (MMQ) (4, pp. 94-100). There was no particular reason for the selection of these tests, except that they were part of a larger test battery designed to investigate leadership qualities. This research was conducted at the Institute for Industrial Supervision in Finland from 1952 to 1954.

Subjects. There was a total of four groups of Ss:

1. One group (S 1) consisted of 102 foremen candidates who were tested with the test battery including the TTS for purposes of evaluating leadership qualities. The Ss knw that about 20 % of them would be selected as factory foremen.

2. A second group (R 1) included 140 foremen from various factories. These Ss were also given a battery of tests including the TTS during the last week of a one month leadership training course.

3. The third group (S 2) included 153 foremen candidates, as S 1, who were given the MMQ in their battery of tests.

4. The fourth group (R 2) included 130 foremen, as R 1, who were given the MMQ in their test battery.

The Ss were instructed to be as unbiased in their responses as they could and to give as accurate a picture of themselves as their self-knowledge permitted. – Scoring of the TTS followed the test manual. In computing the score for the MMQ the 10 least reliable items were excluded (for details see 6, pp. 121-125).

Both candidate groups (S 1 and S 2) knew that the procedure was designed to assist in the selection of new foremen. The foremen (R 1 and R 2) were told that the tests were designed only to investigate personality variables connected with leadership qualities, and that the results would have no bearing on their future. The actual test instructions were the same for all groups. – Thus groups S 1 and S 2 were in a "selection situation" and groups R 1 and R 2 in a "research situation".

As criterion of leadership was used a measure combined from the leadership variable in a standard Leaderless Group Task (LGT) and a sociometric evaluation of leadership in the training courses (6, pp. 95-96 and 103-115, and 1, 2, 5, and 8). The Pearson *r* correlation for these two variables was $.56 \pm .07$. This criterion was available only for the R groups. However, all groups were given a large battery of tests, some of which had a very high correlation with the criterion for the R groups. It seems reasonable to consider scores on these tests as indicators of leadership qualities. Examination of these scores for the S groups reveals no difference between S and R groups. It may therefore be assumed that the S and the R groups did not differ much in real leadership qualities.

RESULTS

The results are presented in detail in Table 1. They show a significant difference between research and selection situation for "Emotional Stability", "Sociability", "Dominance", and "Impulsivity" in TTS and "Neuroticism" in MMQ. There was no significant differences for "Activity", "Vigorousity", and "Reflectivity" in TTS.

The results also suggest, that in the selection situation there is a trend to higher score on those variables shown to have high correlations to the criterion. In other words, it looks as though the Ss in the selection situation "know" which variables are "good" for foremen.

These results make is apparent that motivational conditions may have an important effect on answers to questionnaire type tests. Accordingly, the valid use of such tests would seem to require as neutral a motivational test situation as possible.

Table 1. Variations of the Questionnaire Variables in the Selection and Research Situations and the Correlations to Leadership Criterion.

Variable and group					Mean diff.			Correlation to leadership criterion	
		Q-1	Median	Q-3	M_{S1}-M_{R1}	t	p	r	p
Dominance	R 1	4.7	6.5	9.2	1.77	3.82	.001	.35	.001
	S 1	5.6	8.6	11.0					
Emotional	R 1	7.0	9.4	12.2	3.24	7.36	.001	.26	.01
Stability	S 1	11.1	13.2	15.3					
Vigorousity	R 1	7.8	10.1	12.2	.53	1.35	–	.23	.01
	S 1	8.6	11.0	12.5					
Sociability	R 1	4.8	7.7	10.0	2.18	5.23	.001	.19	.05
	S 1	7.7	9.9	12.1					
Reflectivity	R 1	5.8	8.0	10.6	.70	1.68	–	.19	.05
	S 1	6.5	8.7	10.8					
Activity	R 1	6.7	8.7	11.2	.44	.33	–	.12	–
	S 1	7.0	8.5	11.4					
Impulsivity	R 1	6.8	8.8	11.0	1.11	2.93	.01	.04	–
	S 1	8.0	9.5	11.6					
Neuroticism	R 1	4.0	6.6	9.3	–2.42	5.44	.001	.15*	.001
(MMQ)	S 1	2.5	5.0	6.7				(.46)**	

Note: R 1 and R 2 are groups in a research situation, group S 1 and S 2 groups in a selection situation. Correlations can be computed only for the R-groups.

* $N = 356$.

** Dichotomous variable. The limit of the two classes is + 0.56. $N = 53$.

REFERENCES

1. *Assessment of men.* Edited by the OSS Assessment Staff. New York: Rinehart, 1948.
2. Bass, B. M.: The leaderless group discussion technique. *Personel. Psychol.*, 1950, *3*, 17-32.
3. Edwards, A. L.: *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Eysenck, H. J.: *The scientific study of personality.* London: Routledge & K. Paul, 1952.
5. Harris, H.: *The group approach to leadership testing.* London: McGraw-Hill, 1949.
6. Rainio, K.: Leadership qualities. *Annales Academiae Fennicae.* Helsinki, 1955.
7. Thurstone, L. L.: *Thurstone Temperament Schedule.* Chicago: Univer. of Chicago Press, 1949.
8. Vernon, Ph. E.: *Personality tests and assessment.* London: Methuen, 1953.

A GENETIC APPROACH TO SEX DIFFERENCES IN LOCALIZATION

C. I. SANDSTRÖM and I. LUNDBERG
University of Upsala, Sweden

PROBLEM

It has previously been shown that sex differences occur in adult Ss in opto-kinesthetic (1) and tactual-kinesthetic perception (2). In this paper we shall consider if sex differences can also be observed in children (8-15 years). For this purpose we decided to work extensively with one of the opto-kinesthetic experiments previously used and particularly suitable for relatively young Ss.

In accordance with former definitions by the senior author (1, p. 94) we regard the experiment as a localization test.

PROCEDURE

The apparatus consists of a specially constructed table. The top is an ebonite plate 4 mm thick (750 × 650 mm) which is painted white. There is a black fixation point (diam. 1.5 mm) located medially 240 mm inwards from the long side where S is seated.

While fixating the point binocularly S is instructed to place manually a short mapping pin in the spot under the table-top, which he considers corresponds with the position of the fixation point. The mapping pin is held between forefinger and thumb, and in order to make it easier to stick the pin into the plate, a thin piece of rubber was glued onto the underside of the table-top. S may not use his tactual sense, but must make his determination only by his conception of the position of his hand (body image).

During the experiment E sits directly opposite S. The table-top opens towards S, and thus the results (the points of impact) can be read without this being seen by S. A sheet of squared paper (in mm) is fixed to the underside of the table-top, and the point of origin of the co-ordinate system drawn on this paper coincides with the fixation point. The position of the mapping pin on the x- and y-axes is then marked.

Each S carried out the experiment four times with both his right and his left hand; the right-hand and the left-hand experiments were alternated systematically.

This investigation was supported by a grant from the State Institute of Psychology and Education.

The investigation was conducted with 480 children from the Fredhäll primary school (Stockholm), aged 8-15 years (born 1947-1940). Each age group was represented by 30 Ss of each sex.

In no case were any difficulties experienced in connection with the instructions. E sat at the table and demonstrated how the experiment should be conducted: "You are to fixate this black point on the table-top, and then try to stick such a pin exactly under the black point. But you must continue all the time to look at the fixation point." Ss were well motivated, and E could easily control that they had understood what they had to do.

RESULTS

In reporting the results, we have taken into consideration two types of means, and have designated these, as previously, distance values and centre-of-gravity values.

By distance values we mean the actual distance of the point of impact from the point aimed at, irrespective of direction. (All the x-values, irrespective of their signs, are added together; the same applies to the y-values, and the distance $= \sqrt{x^2 + y^2}$).

The centres of gravity refer to the values obtained when the points of impact on the x- and y-axes are considered in accordance with their plus and minus values. If no factors determining direction are operative, these centres of gravity should naturally coincide with the objective point aimed at. This it not the case.

The results obtained with adults (1) are also given.

Distance values

Among the age groups we find only one sex difference of the same order of magnitude as was observed for adult Ss, namely for the 11 year old, left hand. CR based on SD for the 8 year level is 2.60, and based on SD for the 15 year levels is 1.97. CRs for adults are 3.82 (right hand), and 2.32 (left hand).

A tendency towards greater distance values is, however, observed among the girls – the total mean value, right hand, for boys is 15.2, for girls 15.8; the corresponding values for the left-hand are 14.8 and 16.8.

As in the case of adults, the similarity in the results between experiments conducted with the right and with the left hand is very great. In spite of the experience of greater motor uncertainty and resistance, the performances with the left hand are not inferior. This appears to imply

Table 1 shows distance values (in mm) for the different age groups, and SD for Ss aged 8 and 15 years, as well as for adults. The lower values for female (F) Ss compared with male (M) Ss are italicized.

Age groups	Right hand		Left hand	
	M	F	M	F
8	15.0	17.8	18.3	18.3
(SD)	(7.9)	(11.9)	(10.2)	(9.7)
9	13.8	15.5	13.9	17.2
10	16.8	*15.8*	13.5	16.4
11	12.8	16.2	13.0	18.0
12	17.2	*16.5*	16.1	16.8
13	15.1	16.9	15.6	*15.3*
14	13.1	*12.8*	12.1	16.8
15	18.4	*15.0*	16.3	*15.9*
(SD)	(7.7)	(9.0)	(7.5)	(7.5)
Adults	17.1	27.6	17.9	25.1
(SD)	(7.5)	(11.3)	(10.6)	(10.9)

that the sex differences cannot be explained from some simple motor aspect.

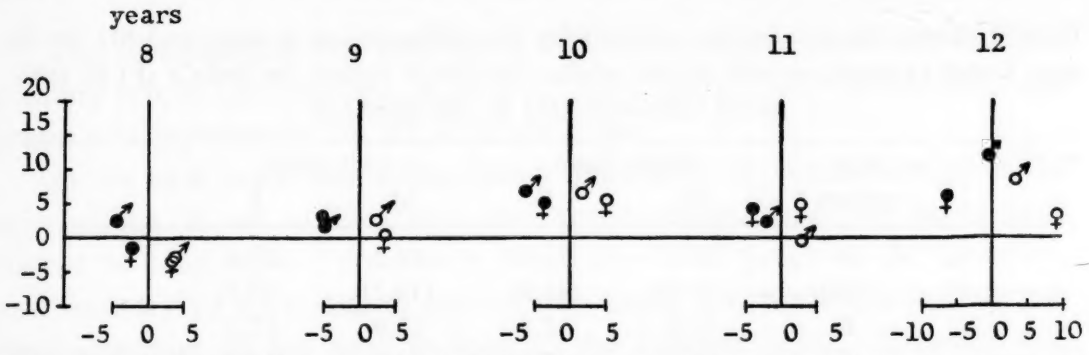
Nor does it seem possible to detect any direct development of the localization ability as measured by this test. The lowest mean value, 12.1 mm, was obtained by boys of 14, left hand, and the highest value, 18.4 mm, by boys of 15, right hand. Half of the 32 mean values for the eight age groups fall within the range of variation 15.0 - 16.9 mm. On the 15 year level the values obtained by the girls are lower than those for the boys both for the right and the left hand.

Centre-of-gravity values

In as far as any sex differences in the distance values are observed, these correspond with the differences found in the adult sample. From the figure it will be seen that as regards the centre-of-gravity values quite a different tendency is found to be dominant.

Sex difference in the adult sample consists in the fact that women are inclined to point further away than men; this difference is definitely significant for both hands. The figure shows, however, that for practically all the age groups, conditions are reversed. Moreover, it should be pointed out, that throughout, the right-hand values fall to the left of the *y*-axes, and that the left hand values fall to the right of this axis, which was not the case in the adult study (cf. Fig.).

The results obtained by Ss 8, 12, and 15 years old have been scrutinized more closely, and their means, deviations, and significant CRs are



The figure (above and to the right) shows the centres of gravity for all age groups as well for adults. The solid sex signs represent right hand values, the other ones left hand. Scale 1:1.5.

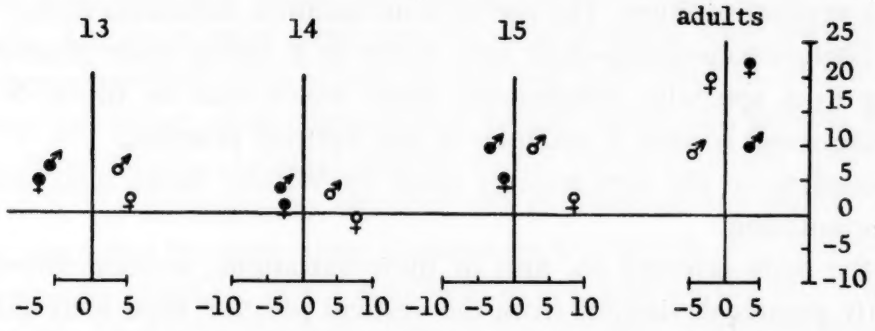
shown together with those for the adult sample in Table 2. The highest CRs are found on the *x*-axis. Only one CR over 2 (2.25) is found on the *y*-axis; namely, for the right hand, 15 year level. This difference is quite contrary to that found in the adult sample.

As in the case of their distance values, Ss 15 years showed also the most pronounced differences in their centre-of-gravity values as compared with the adult sample. The only age group which showed a tendency to correspond with the adult sample was the 11 year old Ss, and as was pointed out above, this applies also to distance values.

The remarkable fact, that the values on the *x*-axis for each hand fall throughout inversely for children as compared with adults, cannot be commented on here, as further data are required.

Table 2 shows centres of gravity, SDs, and significant CRs for Ss aged 8, 12 and 15 years, as well as for adults. Values in mm.

Age groups	Right hand				Left hand			
	x-axis		*y*-axis		*x*-axis		*y*-axis	
	M	F	M	F	M	F	M	F
8	–4.5	–2.2	2.3	–1.4	4.1	3.7	–3.0	–3.2
(SD)	(9.3)	(13.0)	(12.0)	(14.6)	(13.1)	(12.4)	(12.8)	(14.4)
(CR)	—		—		—		—	
12	–0.5	–6.3	11.9	6.1	3.5	9.2	8.8	3.3
(SD)	(7.8)	(7.4)	(12.6)	(13.5)	(7.8)	(10.8)	(13.0)	(13.3)
(CR)	(2.90)		—		(2.31)		—	
15	–3.4	–1.4	9.4	5.3	2.4	8.7	9.4	2.2
(SD)	(9.7)	(7.6)	(14.6)	(14.3)	(8.5)	(9.2)	(11.9)	(12.7)
(CR)	—		—		(2.86)		(2.25)	
Adults	3.8	3.5	10.0	22.0	–4.7	–2.1	8.8	20.3
(SD)	(9.1)	(15.3)	(10.4)	(11.4)	(12.4)	(13.0)	(13.1)	(11.5)
(CR)	—		(3.81)		—		(3.22)	



DISCUSSION

We have found that though the sex differences as regards distance values are small for Ss aged 8-15 years, they tend to correspond with those in the adult sample. It may be asserted that these differences would have proved rather uninteresting had there not existed significant differences at adult level.

As regards the centres-of-gravity, an inverse tendency to that found in the adult sample is observed, and in one case CR is 2.25. This refers to the right-hand values of Ss aged 15 years, and we would emphasize that, in respect to sex differences, this age group deviates most from the adults.

Practically in all cases children obtain better results than adults, and the difference between adult and juvenile female Ss is highly significant. The mechanisms operative in this type of position determinations appears to deteriorate as regards precision during later adolescence.

It may be objected that the present sample consists of a more representative average of the population than the sample comprising university students which was employed for investigating adults. If the investigation of adults on the basis of a wider section of the population did not, however, show any sex difference, we should be faced by the curious question why precisely female, as compared with male, students should possess an inferior equipment in respect of the functions measured by the test.

Witkin, to whose investigations we have referred in earlier papers, has applied his tests also to children and adolescents of different ages (8, 10, 13, 15, and 17 years of age). The results which he obtained were remarkably similar to those reported here. This applies to his rod-and-frame test, tilting-room-tilting-chair test, and embedded figures test. For a description and report on these tests the reader may consult Witkin's comprehensive and detailed work "Personality through perception" (3).

Especially the first two tests mentioned appear to be relevant. In the rod-and-frame test, S sits on a chair which is either in a vertical or a given inclined position, and has to adjust a pivoted luminous rod so that it oc-

cupies a vertical position. The rod is in an inclined, luminous square frame. In the tilting-room-tilting-chair test, S sits in a tilting chair placed in the opening to a specially constructed room which can be tilted. S has to adjust the room so that it occupies a true vertical position.

According to the terminology used by Witkin, these tests deal with space orientation.

In the tests referred to, and in their variations, women showed significantly greater deviations from the vertical position than men did – i. e., the rod or the room was actually on the average in a more inclined position when judged to be vertical by the women than by the men. This sex difference was apparent, though less marked, for juvenile age levels, but "despite this general trend, it is at the same time apparent that the differences between males and females at the 8-, 10-, 13-, and 15-year levels are for the most part slight. Males and females appear to be most nearly alike in perception at *the 15-year level, where the differences in mean scores tend to be exceedingly small.* – It is only at the adult level, however, that a picture of consistently significant sex differences emerges." (3, p. 147 f. Our italics.)

Witkin shows, as do also our results, that younger Ss often have smaller deviations than adults, and this is particularly pronounced in the case of girls.

This agreement with Witkin's results would seem to make it reasonable to assume that similar functions are decisive for the results, and that the same explanation of sex differences at adult level holds good for both his and our tests. The clue, which perhaps especially the remarkable concordance at the 15-year level appears to give, is, however, not easy to make use of.

In his attempt to explain the sex differences. Witkin starts by assuming that men and women are dependent in different ways on the external (e. g., the luminous frame and rod, the specially constructed room) and internal (body image) stimulus field. Women are liable to accept the external field in a more passive manner, while men are inclined to deal with this field more actively and analytically. Men are more concerned with sensations from their bodies, whereas women pay less attention to stimuli arising from the body.

This is plausible and is supported also by behavior observations. Sandström's adult female Ss participated more spontaneously in the experimental situation, they accepted the instructions more unreflectingly, and, in general, adopted a more passive attitude to the task. As a rule, the male Ss were more suspicious and wondering, and usually required more time to make their decisions.

But what has now been said about adult behavior, applies also to young Ss. The boys were often very inquisitive about what was under the table; they wanted to know where the pin had been stuck in, and wondered what was the object of the experiment. In the higher age groups, the boys frequently adopted a sceptical attitude to the whole situation, and some of them tried, for example, to estimate the distance of the point from the edge of the table by using their forearm as a measure. All such impulses were naturally checked, but it is typical that E had hardly any difficulties at all of this kind with the girls. Very little curiosity was observed in their case, and they seldom asked any question about the purpose of the experiment.

A difference in behavior between the adult and the juvenile male Ss would appear to be descernible: the younger boys at any rate displayed great eagerness in undertaking the task, as if they were seized by some technical fascination inherent in the situation, while in the adults this eagerness was replaced by a sceptical and more cautious attitude.

What has been pointed out with regard to behavior, does not appear to give the slightest clue to the elucidation of the question as to sex difference, and in this connection appears to somewhat weaken Witkin's assumption. In any case, these facts do not seem to deal with essentials. The eager boys, who really do seem to surrender themselves to the external stimulus field should unquestionably obtain worse average results than adult male Ss. Why girls achieve better results than women remains also entirely unexplained.

It was also pointed out previously (2), that the intraindividual consistency of the performances does not apparently tally with Witkin's hypothesis.

REFERENCES

1. Sandström, C. I.: Sex differences in localization and orientation. *Acta psychol.*, 1953, *9*, 82-96.
2. – Sex differences in tactual-kinaesthetic and visual perception of verticality. *Quart. J. exp. Psychol.*, 1956, *8*, 1-7.
3. Witkin, H. A., *et al.*: Personality through perception. N.Y.: Harper, 1954.

CONCEPTS IN PSYCHOLOGICAL MODEL CONSTRUCTION

PER SAUGSTAD
University of Oslo, Norway

In a recent article in this journal Rommetveit (24) has discussed certain problems relevant to the use of concepts in theoretical models in psychology. In reading his defense of "surplus meaning" the first impression was that it raised some interesting questions. A critical examination of the article has, however, led me to doubt that his assertions are sufficiently well formulated to serve as a fruitful basis for further discussion. In two respects I consider his presentation unfortunate and would like to submit some critical comments:

I. The article does not give an adequate presentation of the points of view expressed by other writers.

II. The presentation tends to oversimplify the complexity of the problems involved.

I

The phrase "surplus meaning" was apparently coined by Reichenbach (23), but the problems involved are old in philosophy. The writings of the positivists brought the question of whether or not scientific concepts might be reduced entirely to an observational basis to the foreground. This problem was made the subject of penetrating analysis by members of the Vienna Circle. Carnap's (5) attack on the problem by his introduction of the so-called "reduction sentence" is well known. In a recent monograph Hempel (9) has given what to me seems a balanced, systematic, and sophisticated analysis of problems arising in connection with the construction of scientific concepts. This monograph devotes considerable space to the use of concepts in psychological research. Hempel strongly emphasizes that to be of significance scientific concepts must have what he refers to as "systematic import" and "theoretical import". In the construction of scientific concepts, neither the theoretical significance nor the observational basis must be neglected. Like Rommetveit, Hempel also discusses problems relevant to the use of concepts taken over from everyday knowledge and thus couched in conversational language. As might be expected the discussions of the philosophers have had wide repercussions among psychologists, and an extensive literature relevant to problems concerning the use of concepts in psychological models has accrued. If not in the same clear form, many

of the arguments given by Hempel have been presented. The list below contains some of the better known names: Bergman and Spence (2), Bohring (3), Carnap (5, 6), Feigl (7), Hempel (9), Koch (13), MacCorquodale and Meehl (15), Marx (17), Pratt (21, 22), Reichenbach (23), Stevens (26).

Even if I do not understand Rommetveit's argumentation sufficiently well to feel I might fruitfully engage in a discussion of whether or not it might be wise for a research worker to follow his instructions, I consider it evident that other writers have dealt with the problems from much the same point of view. If Rommetveit has added anything important to the discussion of these central problems of methodology, it seems reasonable that he makes his claim valid by relating his position in some detail to that of other writers. If the article is not written with the intention of adding anything new to the discussion, it is unfortunate that no reference is made to writers having dealt with the same aspects of the problem, as much may be gained from a study of the literature concerned.

Rommetveit has referred to other writers, but only to those expressing ideas which are supposed to differ from his own in important respects. No reference is made to writers expressing more or less the same points of view as himself this being the occasion of the complaint expressed above. However, with regard to the five articles referred to as representative of viewpoints differing from his own, I also have some comments to make. As will be understood, I am not able to decide whether the points of view offered by these writers actually differ from his, but I want to make clear that the articles have not been adequately characterized by Rommetveit in all cases. In a passage where reference to the five writers is made, one finds the following statement: "The demand for unequivocal empirical anchoring is then set up as a superordinate norm for psychological research: A concept, it is maintained, acquires its entire meaning from the rules coordinating it to observations" (p. 339). I will deal with the five articles one by one from the point of view of whether or not they conform to the characterization given in this excerpt, which, by the way, seems to be in line with the other assertions in the passage.

1. Ginsberg (8) has made an analysis of the terms "intervening variable" and "hypothetical construct" as introduced by MacCorquodale and Meehl (15). He offers the interesting point of view that a distinction between these terms might be relatable to a distinction between a scientific law and a scientific theory. Considering Ginsberg's notions of what is characteristic of a theory, it will be clear that his position can not be the one attributed to him by Rommetveit. Ginsberg states: "In contrast to a law, a theory is not a product of inductive generalization nor is it con-

stituted by concepts which are explicable in observation terms alone. Theories are *invented,* highly abstract and general systems of concepts whose association with observations or facts is indirect rather than explicit. No amount of data inevitably determines a theory in the same way as it determines a law." (P. 121). In many respects Ginsberg's point of view implies a criticism of the ideas of MacCorquodale and Meehl. Ginsberg also makes explicit a point which Marx and, as far as I can see, Rommetveit have failed to make. Since "surplus meaning" is characteristic of concepts even in the most highly developed scientific theories, it is thus not something that must be tolerated only at the present stages in the development of psychological theories. It is something essential to any good scientific theory. A number of passages might be quoted as indications of the incorrectness of Rommetveit's presentation of Ginsberg's point of view. One more may suffice. After having stressed what he calls the systematic purport of a useful scientific concept, he writes: "This openness of meaning implies that there is always an indeterminate range of factual content implicit in any significant concept" (p. 127).

2. Kessen and Kimble (12) discuss certain uses of neurological concepts in psychology. At the very end of their paper brief mention is made of what they consider reasonable requirements for a scientific theory. Complaint is here made regarding a definite model that it does not contain a single operational definition. As long as the writers have not specified what they consider is an operational definition, it is unjustifiable to attribute to them the point of view indicated by Rommetveit in the above statement. Actually the last two pages of their article contain passages that make it fairly evident that they do not conform to Rommetveit's characterization.

3. Maatsch and Behan (14), I admit, have given a presentation of requirements to a theoretical language apparently conforming to Rommetveit's characterization.

4. Marx (16) uses the terms "intervening variable" and "hypothetical construct" in a sense somewhat different from that suggested by MacCorquodale and Meehl: "By intervening variable is meant any intervening construct with a maximum amount of operational validity, or direct empirical reference, and by hypothetical construct is meant any construct with a relatively low degree of operational validity" (p. 114). From the point of view that psychological concepts very often tend to have a rather unspecified observational basis, one may apparently agree with Marx' statement that the ultimate objective of all theoretical construct formation is this type of intervening variable, and still maintain that it is neither possible nor desirable to free psychological concepts of their "surplus mean-

ing". I am unable to see that the acceptance of this objective implies setting up unequivocal empirical anchoring as a superordinate rule. It should also be noted that Marx, in prescribing rules for the transformation of "hypothetical constructs" into "intervening variables", is very cautious.

5. Spence's article (25) contains some sentences that may qualify him for Rommetveit's characterization. On the other hand, passages are found suggesting that his position is not easily described, e. g. when he argues for a phenomenological approach under certain circumstances (p. 77).

This brief review suggests that Rommetveit's characterization can not reasonably be considered definitely correct for more than one, possibly for two, of the five writers, that it is unjust for one, and that it is not applicable at all to the remaining two. This oversimplification is unfortunate, as it may give the impression that views set forth by the methodologists are far more uniform that they actually are.

In connection with Rommetveit's characterization of other writers, I will only briefly comment on the following passage from his article: "Indeed, by tacitly assuming a strange dichotomy between a region of pre-scientific revelations and a realm of "scientific" (i. e. operationally defined) terms, the preachers of operationalism ultimately stand out as the mystics of modern scientific psychology" (p. 344). As has repeatedly been pointed out, it is difficult to delimit the meaning of the term operation (Ness (20), Stevens (26), Israel (11)). It may, therefore, not be simple to make clear who are to be considered preachers of operationism. From what is gathered from the previous discussion in Rommetveit's paper, Marx ought to qualify as one of the "mystics". In the article referred to by Rommetveit Marx has made the following statement: "From an historical point of view, logical constructs, like scientific problems and hypotheses, originate in the common-sense reflections of men". Anyone familiar with Bridgman's writings will remember his emphasising that explanation consists in reducing the unfamiliar to the familiar (cf. 4). It might also be noted that one of the most eager proponents of operationism in psychology has been Tolman, whose writings give little evidence to show that he conforms to Rommetveit's characterization above.

II

Analyses by e. g. Carnap (5) and Hempel (9) make one doubt that all scientific concepts may be reduced entirely to observations. In his monograph Hempel argued convincingly that neither quantitative nor the so-called disposition terms can be reduced to what he calls "observables". Every scientific theory will thus necessarily contain some concepts with a

"surplus meaning". On p. 343 Rommetveit writes: "If this residual of as yet inexplicitly stated but vaguely anticipated semantic rules is eliminated by a single operational definition (in part dictated by narrowly pragmatic considerations concerning available research instruments) the researcher will sever his contact with the universe of extra-laboratory observations in a most unfortunate way." This statement may give the impression that it is possible, if the research worker so desires, to anchor terms unequivocally to an observational basis. Whatever Rommetveit's intentions have been here I think he should have made the above point explicit. Before one proceeds to discuss whether "surplus meaning" is desirable or not, it seems reasonable to ask first whether "surplus meaning" can be avoided in scientific theory construction.

The following passage is apparently central to Rommetveit's argument: "The fact is that once we begin analyzing the meaning of a given concept we are forced to recognize that it does not occur in a theoretical vacuum, but forms part of a more inclusive model. It therefore acquires its meaning from at least two sources, from *1) the semantic rules and 2) the syntactic rules coordinating the concept to other concepts within the more inclusive research model.*" (P. 339). Concerning the "syntactic rules" Rommetveit states previously: "The *syntactic rules* are then stated or implicit logical relations existing between concepts within the model and the *semantic rules* are the rules set up for anchoring concepts empirically" (p. 336). Now if the "syntactic rules" are implicit it is apparently very difficult to attach meaning to a term by means of them. In attempting to construct a model, the research worker in psychology will be placed in a situation where his "syntactic rules" are to a considerable extent unstated. With this consideration in mind, it may be difficult to understand what Rommetveit has intended by the above statement. If Rommetveit has meant to say that the research worker also has to consider that in order to be scientifically fruitful his concepts will have to enter into a logical relationship to other concepts, in addition to having an empirical meaning, this seems to be a statement of the obvious, as the aim of science for a research worker is the construction of a theory. Likewise, if it is meant that the research worker ought to utilize the implicit logical relations between the terms from everyday language, the statement is understandable to me, as apparently this is what every research worker in psychology necessarily does in attempting to construct fruitful concepts.

In connection with the unstated "syntactic rules" it might be mentioned that Rommetveit may have misunderstood an important point concerning "the intervening variable" as introduced by Tolman (27) and Hull (10).

In section 2 Rommetveit has discussed the "intervening variable" and has made clear that it is anchored on two sides as in the example from Tolman where "the intervening variable" is determined in relation both to time since last feeding and to motor learning. One may now ask why Tolman did not determine his "intervening variable" by time since last feeding and e. g. the number of hairs in the rat's fur. The answer evidently is that Tolman has considered some implicit "syntactic rules". The use of the "intervening variable" thus presupposes both "semantic rules" and "syntactic rules". On p. 339, where the above quotation is found, one gains the impression that methodologists in discussing the "intervening variable" have been apt to overlook what Rommetveit calls the "syntactic rules".

The following passage from Rommetveit is also relevant to the statement about the tacit "syntactic rules" and likewise important in his argumentation: "As a general norm for research, then, juxtaposed to the demand for unequivocal empirical anchoring, we assert: We should construct our research models so as to provide for optimal use of prescientific knowledge" (p. 344). Again I am unable to see what this assertion amounts to. Starting out to construct a model for some unexplored area, the research worker will find himself in the following situation: neither the "semantic rules" nor the "syntactic rules" are stated. Further, his choice of "semantic rules" will influence his choice of "syntactic rules" and vice versa. To be advised in such a situation to make optimal use of prescientific knowledge is evidently of little help. He might probably just as well be given the advice to do as well as he can. What was optimal would have to be determined in relation to some good theory, but in this case, a good theory is not in existence. Before Rommetveit has indicated what an optimal use of prescientific knowledge may amount to, I can not see how one could agree or disagree to his assertion.

It may be difficult to conceive of what considerations led Rommetveit to make the above assertion. However, it should be pointed out that the way he introduces his terms "syntactic rules" and "semantic rules" may be highly misleading. In sketching what he considers important aspects of a theoretical model relevant to the discussion in his article, Rommetveit avails himself of the analogy of the geographical map, and thus compares the task of the model-builder to that of the cartographer. In using this old analogy Rommetveit might have referred to other writers. The advantages and disadvantages of this analogy have recently been discussed by Toulmin (28), who points out that in one respect the analogy is a most unhappy one. In the case of the geographical map, what Rommetveit has called the "syntactic rules" are determined in advance before the carto-

grapher starts out to perform his task. For the research worker, on the other hand, the important relations between his concepts are to be discovered. The analogy thus leaves no room for discovery and inventiveness. Rommetveit himself has stated the analogy is over simple and would probably agree to this comment. Nevertheless, this point is so important that reference to it should not be omitted.

Apparently Rommetveit's argumentation pivots to a considerable extent around the terms "syntactic rules" and "semantic rules". These terms, therefore, ought to have been clarified in connection with the introduction of the research model. This might have made it easier to understand his argumentation and would have made it possible to decide whether he is saying anything additional to what every research worker in psychology necessarily must know. Morris (19) has discussed the terms "syntactics" and "semantics". One disadvantage of the terms is that one might be led to overlook that the "syntactic" and "semantic" aspects in psychological research are necessarily highly interrelated. In his comments to the "intervening variable", Rommetveit seems to have lost sight of this close interrelatedness. Furthermore, the term "rule" can be misleading. By a rule one does usually conceive of somewhat clear and unambiguous relations. When the implicit logical relations between terms from conversational language are considered, they in many cases appear to be extremely vague and often dealt with on an intuitive basis.

In connection with the role of discovery in stating the "syntactic rules" a remark should be made to some sentences by Rommetveit concerning the role of observation in research. On p. 344 he writes: "Suppose we happen to find a psychologist who outside the laboratory has made observations that may in the future prove as crucial to the development of psychology as those made by Newton in the orchard turned out to be to natural science." Whatever the basis for the story about Newton and the apple (see Andrade (1)), it must be assumed that millions of people before Newton have seen an apple falling. There is nothing great or ingenious in this observation. What would be the result of the work of a genius, of course, is the system in which this observation is correlated to other events. On an earlier page the same undue emphasis on the role of observation seems to be expressed. "But unlike most cartographers the research psychologist frequently finds himself standing on a most meagre and insufficient observational basis." This again indicates that the analogy of the geographical map may lead one astray. The most striking difference between the cartographer and the research worker in psychology apparently is that the psychologist has no good system which he can use for ordering his data. It

may further be considered doubtful whether his observational material is meagre. On the contrary, I personally believe that one of the main difficulties in research in psychology is due to the richness of available observations.

The question of "surplus meaning" constitutes a number of intricate problems for psychologists and it is very difficult to present a balanced account of the issue. Without the historical background, an account will necessarily be distorted. The history of psychology is replete with instances where one research worker or a group of research workers has been unable to communicate to colleagues in an efficient way. The struggle over psychoanalysis and "imageless thoughts" are pertinent cases. The tendencies towards a more rigourous theoretical language as instanced by behaviorism, operationism and logical positivism in psychology, must be viewed against this background. In a few introductory sentences Rommetveit has referred to this difficulty (p. 336). This seems to be a somewhat superficial treatment of an embarrassingly complicated problem.

Further, in presenting what he considers an appropriate type of research model, one meets with the same superficial treatment of the problem of communication. Rommetveit intends to show the reader what might be gained by the use of a more flexible research model, and argues for what he later on refers to as partial delimitation of model signs. This procedure he claims may lead to any one of the following types of revisions: 1. A revision of semantic rules. 2. A delimitation of field of predication. 3. A revision of the model syntax. The construction of a model capable of these revisions constitutes, I believe, what most research workers in psychology would have as their aim; at least they would nourish it as an ideal. As the history of psychology shows, the construction of models capable of being revised in one of the three ways suggested involves great difficulties. Rommetveit's claim that his model is stated in this way thus seems to be a very presumptuous one. I am unable to see what type of observation would make it clear that his model needed one of the three types of revisions. Rommetveit's attempt to achieve this in one page convinces me that he has not paid sufficient attention to the difficulties of empirical research in psychology.

SUMMARY

Some critical comments are made to a recent article in this journal by Rommetveit who has dealt with problems relevant to the use of concepts in psychological models. It is maintained: I. that Rommetveit has given an inadequate presentation of the points of view of other writers; and II. that his presentation tends to oversimplify the complexity of the problems involved.

REFERENCES

1. Andrade, E. N. da C.: Newton and the science of his age. *Proc. Roy. Soc. Lond.*, 1943, *B 131*, 207-223.
2. Bergman, G., & Spence, K. W.: Operationism and theory construction. In (18), 54-66.
3. Bohring, E. G.: Symposium on operationism. The use of operational definitions in science. *Psychol. Rev.*, 1945, *52*, 243-245.
4. Bridgman, P. W.: *The logic of modern physics.* New York: MacMillan, 1927.
5. Carnap, R.: Testability and meaning. *Phil. Sci.*, 1936, *3*, 419-471; 1937, *4*, 1-40.
6. – *Logical foundations of probability.* London: Routledge and Kegan Paul, 1950.
7. Feigl, H.: Symposium on operationism. Operationism and scientific method. *Psychol. Rev.*, 1945, *52*, 250-259.
8. Ginsberg, A.: Hypothetical constructs and intervening variables. *Psychol. Rev.*, 1954, *61*, 119-131.
9. Hempel, C. G.: *Fundamentals of concept formation in empirical science.* Int. Encyc. unif. Sci., II, 7, 1952.
10. Hull, C. L.: *Principles of behavior.* New York: Appleton-Century, 1943.
11. Israel, H. E.: Symposium on operationism. Two difficulties in operational thinking. *Psychol. Rev.*, 1945, *52*, 260-261.
12. Kessen, W., & Kimble, G.: "Dynamic systems" and theory construction. *Psychol. Rev.*, 1952, *59*, 263-267.
13. Koch, S.: The logical character of the motivation concept. *Psychol. Rev.*, 1941, *48*, 15-38; 127-154.
14. Maatsch, J. L., & Behan, R. A.: A more rigorous theoretical language. *Psychol. Rev.*, 1953, *60*, 189-196.
15. MacCorquodale, K., & Meehl, P. E.: Operational validity of intervening constructs. In (18), 103-111.
16. Marx, M. H.: Hypothesis and construct. In (18), 112-129.
17. – The general nature of theory construction. In (18), 4-19.
18. – (ed.): *Psychological theory. Contemporary readings.* New York: MacMillan, 1951.
19. Morris, C. W.: *Foundations of the theory of signs.* Int. Encyc. unif. Sci., I, 2, 1938.
20. Ness, A.: *Erkenntnis und wissenschaftliches Verhalten.* Oslo: Dybwad, 1936.
21. Pratt, C. C.: *The logic of modern psychology.* New York: MacMillan, 1939.
22. – Symposium on operationism. Operationism in psychology. *Psychol. Rev.*, 1945, *52*, 262-269.
23. Reichenbach, H.: *Experience and prediction.* Univ. Chicago Press, 1938.
24. Rommetveit, R.: Model construction in psychology: a defence of "surplus meanings" of psychological concepts. *Acta psychol.*, 1955, *11*, 335-345 (& *Nord. Psykol.*, 1955, *7*, 159-169).
25. Spence, K. W.: Types of constructs in psychology. In (18), 68-86.
26. Stevens, S. S.: Psychology and the science of science. In (18), 21-54.
27. Tolman, E. C.: The intervening variable. In (18), 87-102.
28. Toulmin, S.: *The philosophy of Science.* London: Hutchinson's Univ. Library, 1953.

STUDIES IN THE DEVELOPMENT OF A PERCEPT WITHIN VARIOUS CONTEXTS OF PERCEIVED REALITY

GUDMUND SMITH and MAJ HENRIKSSON
University of Lund, Sweden

In a previous paper (10) the present authors described a new tachistoscopic technique for studying perceptual processes beyond awareness, and notably the development of such processes. Two stimuli, A and B, were presented in rapid succession. While the exposure time for B was kept constant at a threshold level where B could be correctly perceived by S, the exposure time for A was gradually increased from sub-threshold values until A emerged as a conscious percept. Experiments were arranged so as to render possible quantitative statistical treatment of data. Interest was centered around that portion of the series where S still did not see A. But there could be no doubt that in a majority of cases the perception of B was affected by the perception of A even before the latter percept had reached the threshold of awareness.

It has been demonstrated before, of course, that the quality of awareness is not indispensable for a definition of perception (for literature see reference 10). But our observations of these pre-stages of conscious percepts registered as change in B convinced us that they should not be regarded only as experimental curiosities, but rather as findings fundamental for explaining the origins of ordinary perception. This second conclusion should be evaluated on the background of our main working hypothesis (4, 6, 8, 10) namely that every objective perception, i. e., a perception more or less coinciding with stimulus reality, may be regarded as the result of a momentary process of organization, an "actual genesis" partly re-constructed by means of the technique described and similar methods (cf. 4, 5, and below). Our interest is therefore not primarily centered around the more or less established products of perceptual processes, but around adaptation processes leading up to ordinary perception and mechanisms regulating the adaptation.

The present authors are indebted to psychiatrists at the Psychiatric Clinic and the St. Lars Mental Hospital in Lund for their valuable co-operation, and to Magnus Bergvalls Stiftelse and Personaladministrativa Rådet for grants supporting this study.

A PILOT STUDY

Before formulating the problem of this investigation the present authors want to report in brief the results of a pilot study. We set out to explore how the perceptual pre-stages observed in our previous series would emerge as independent percepts above the threshold of awareness in various contexts of perceived reality, and also, in case these first independent percepts still belonged to a pre-stage level, their further development in the direction of an objective, established perception. Since the pre-stages of A had to be re-constructed from the established end-stage level, their conscious representation ought to be more or less affected by the perceived context of B. But we also expected pre-stages of A beyond awareness to influence the appearance of B. The genesis of A from initial pre-stages to the established end-stage might probably show obvious inter-individual differences; and part of our interest concerned the stability of some outstanding formal characteristics of these sequences.

The experimental procedure differed from that described below mainly in that exposure times were increased in arithmetical progression. There were six different series. Stimuli consisted of concrete drawings. Since we wanted to study how percepts referring to the same A-stimulus develop within the contexts of contrasting situations (B-stimuli), we chose to select one pair of B-drawings for each single A, one B within the pair being congruent with A (series 1 and 2), the other B being incongruent (series 3 and 4). In order further to highlight our basic problem, we introduced a composite A-stimulus made up of B-stimuli 1 and 3, in the last two series. In series 5 the B-stimulus was the A-stimulus of series 1 and 3; in series 6 a geometrical figure of the same general shape was introduced as a control. Stimuli belonging to series 1 and 3 are presented in Fig:s 1-3. In series 2, B pictured a family scene: a pram, a young woman in a recumbent position and a sitting child, arranged around a blanket. Stimulus A, a naked baby of about 6 months, was going to appear in the middle of this group. The counterpart stimulus B in series 4 showed a couch with three rats, two sitting on the top of it and the third climbing the farther side. All the animals were as big as the human figures in the family group. The baby was exposed between their snouts.

Using selected criteria of correct perception and by ensuring that Ss felt convinced about their interpretation, we were able to bring the perception of B up to an end-stage level before A was introduced (see further below). Furthermore, it proved fairly simple to determine when the first independent representation of A appeared above the threshold of aware-

ness, mainly because stimulus A was projected on an empty space in B (around B in series 5 and 6). According to the main hypothesis, the first representation of A was more likely to be correct in congruent series than in incongruent ones. But we expected in addition that in a series where a stimulus A was exposed for the first time, its correct interpretation would be more difficult than in the next series including the same A. The first four series were thus arranged so as not to promote the hypothesis that the perception of A is going to be partly organized on the pattern of the established perception of B. Similarly, the order of presentation of series 5 and 6 did not favor the expectation that the street-scene parts were likely to be reported more often in series 5 than in series 6.

Altogether 40 Ss (18 men and 22 women) took part in these experiments. All of them were graduate students at the university, their average age being 25.8 years. However, the procedure was gradually elaborated during the first two weeks of testing to ensure stimulus equivalence in all series. As a consequence, only 38 Ss took part in the correct version of the car experiments (described below), 33 Ss in an identical version of the baby experiments, 32 Ss in both of these pairs of series, and 22 Ss in the composite picture experiments.

We restrict our report to some few χ^2-values (always corrected for continuity). The first independent A was more often ocrrect (a car) in the (congruent) street-scene series than in the interior series ($\chi^2 = 12.4$, $P<.001$). The next pair of stimuli offered similar results ($\chi^2 = 10.2$, $P<.01$). In the composite picture series we registered the first object interpretations, not the first reports of "something in the background." There were three kinds of first interpretations: a street-scene or parts of it, an interior or parts of it, the street-scene plus the interior (as, e. g., "a room looking out over a street"). Even counting the latter rare answers in the "opposite" category we found more street-scene answers in the car series than in the geometrical figure series ($\chi^2 = 5.9$, $P<.02$). Al these results confirmed our exceptation that the stabilized perception of B was often decisive for the first conceptualization of A. We may add that among the protocolled "incorrect" reports on A in series 1-4 about 20 % were real interpretations more or less fitting the pattern of stimulus B, as, e. g., rugs on the floor in the interior scene.

But pre-stages of A beyond awareness should also influence the perception of B. We payed attention only to such obvious changes in reports from one exposure to another in series 1-4 as did not show up in the introductory control series (where B was presented alone five times at the level of correct perception). Tendencies to perceive B as a double-exposure

and similar reactions probably denoted that a pre-stage of A was reconstructed solely on the pattern of B. But there was also obvious uncertainty in some Ss concerning the identity of succeeding B's. In other instances B was reported as suddenly becoming darker, more diffuse, covered with haze, etc. (still only toward the end of a complete series when S should really know B). These reactions, which were similar to those reported by Bruner and Postman (1), were attributed to a possible conflict between the restraining influence of the accepted B and the increasing tendency of the A-process to represent itself as a percept, resulting an experience of bewilderment, or, more concretely, of impairment of a perception earlier accepted as established reality. The last, and more common, category of change observed in B concerned such changes in perspective, size, general structure, etc., as often tempted the Es to the teleological conclusion that B changed in order to become prepared for the emergence of A as an independent percept. In the rat series, for instance, the animals were often withdrawn from the spot where the baby was going to be placed at the next exposure.

Our sequence analysis involved the genesis of stimulus A from its first appearance either as a conscious percept or as an obvious influence on B. The categorization concerned continuity and discontinuity in this development and also a common mechanism termed cleavage. Continuous development implied increments in the degree of clarity and correctness (or the sudden appearance of a final A). Discontinuity was consequently characterized by S disqualifying one or more of his own interpretations, retreats to previous phases, etc. We dealt furthermore with such cleavages between an early stage of A and the end-stage of B as *directly preceded* the first objective perception of A, i. e., the correct A did not appear until A was seen without the background B. This temporal or spatial cleavage may be a pre-requisite to the acceptance of A since it was used most often in incongruent series and, as we shall see, was typical of a limited number of Ss. Thus on comparing series 1 and 3 on the one hand with series 2 and 4 on the other we obtained a χ^2 of 7.1 ($P<.01$) with respect to the occurrence of cleavage in both pairs of series. A similar calculation regarding discontinuity yielded a χ^2 of 5.7 ($P<.02$). Our results seemed to indicate that these micro-genetic patterns reflected stabilized mechanisms of cognitive regulation and control (cf. 11).

PROBLEM

In order to elucidate the significance of the new technique for studies in personality as well as to explore the meaning of various response tendencies revealed by it, some of the above experiments together with a "square test" described earlier (10), all of them partly revised, have been applied to groups of psychiatric and normal subjects. If the genetic characteristics measured by the method of successive presentation reflect established forms of reality adaptation they should, in some way or other, make it possible to differentiate between homogeneous groups of various psychiatric subjects in a manner understandable in terms of the specific symptoms of each group as far as these can be interpreted within the operational frame of reference of the technique applied. The present authors intend to return to some aspects of this frame of reference later in the text, but refer to previous publications for a more detailed review (4, 6, 7, 8, 9, 10, 11).

Kragh (4) has discussed in detail whether the developmental process under scrutiny, though an experimental construction, may be regarded as a prolongation of the momentary process of organization supposed to precede everyday perception. On the whole answering the question in the affirmative, he points out that everyday perceptual organization is to a considerable extent characterized by intermittent and subliminal stimulation. Further elucidation of this problem is needed, but is beyond the scope of our exploration. At the present stage of experimental psychology it is, however, no longer possible, to conceive of percepts as instantaneous events uninfluenced by preparatory processes. Any attempts to disclose the prerequisites of ordinary perception must therefore be important, even if the methods available are still partly artificial and insufficently developed.

METHOD

The electronic tachistoscope allowing presentation of the two stimuli immediately after each other on a milk-glass screen has been described before (10). The size of the projected pictures was about 40×40 cm. S was placed in an easy-chair at a fixed distance (about 1.5 m) from the screen. A spot of light at eye-level oriented him toward the middle of the picture to be presented; this spot was so faint that it disappeared when a picture was flashed onto the screen. The room was lighted by a table-lamp behind the controlling-device, the empty exposure-field being consequently brilliant enough to prevent light-shocks. In order to avoid short-circuiting E had to refrain from using the short exposure times employed earlier (from

.003 seconds for A); neither could he expose stimuli continuously, but had to let 60 seconds pass between presentations (a pause most often filled by protocollation). A recent calibration revealed that the time error was less than 8 %.

Three of our previous series were used here. In the first one A is a car (Fig. 1) and B the front of an apartment-building, other buildings, trees, etc., being arranged around it (Fig. 2). Toward the lower middle part of B there is an empty space for A. The second series consists of a fan-shaped pattern of lines (A) and a square (B), the latter seen as a rhomboid when presented together with the lines (10). In the third series, A is the same as in series 1, but B is the interior of a room (Fig. 3). As in the house-picture, there is an empty area for A in the lower middle part of the interior picture.[1] Criteria for the objective perception of the Bs were the same as in the studies mentioned above. In the first B S should place the buildings correctly and report at least two more details and in the third series he should apprehend the room and describe not less than the two main objects (easy-chair and table). Finally, he was supposed to see B in series 2 as a real square.

E started by describing the use of the fixation spot and then presented a picture not belonging to the main series. S was told to "report everything he saw flashed on the screen". The first B was then exposed for .05 seconds. In cases of incorrect or incomplete description B was presented again after one minute for an exposure time 25 % longer than the first, and so on with 25 % increments of exposure times until S had fulfilled the criteria of objective perception. Thereafter the exposure time was reduced to .05 seconds; this time was used for five further presentations of B alone. Next A was exposed for .01 seconds, immediately followed by B (exposed for .05 seconds).[2] The exposure of A was gradually increased by steps of 25 % until S had seen the correct A and B together (or after each other) three times (cf. 12). Pauses between presentations were 60 seconds throughout the experiment. Both picture series proceeded in exactly the same manner. The "square" series should be regarded here only as a necessary interlude between the picture experiments.[3]

This program functioned perfectly in the two picture series. The perception of B was established to such a degree that only an insignificant

1. The luminance of these stimuli measured *at the screen surface* was as follows. Buildings = 1.8 L, Interior = 1.8 L, Car = 1.5 L, Square = 1.8 L, Lines = .9 L.
2. In this way B served to extinguish after-images of A, the duration of which was thereby better controlled.
3. Because of the increased exposure times A was recognized too early by most Ss.

Fig. 1. The A-stimulus in the first and third series.

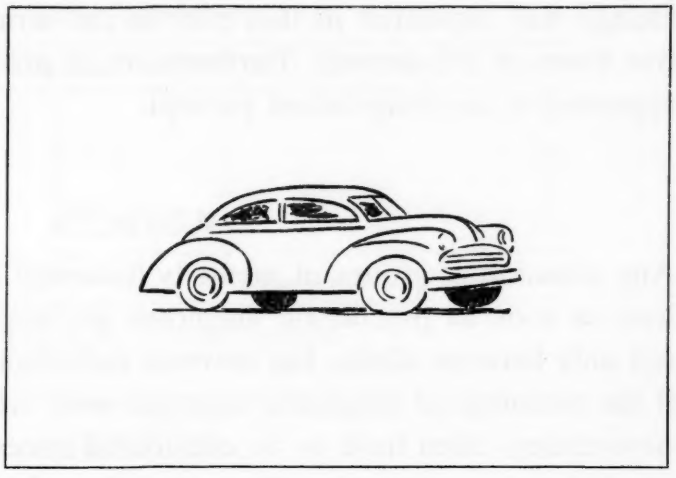


Fig. 2. The B-stimulus in the first series.

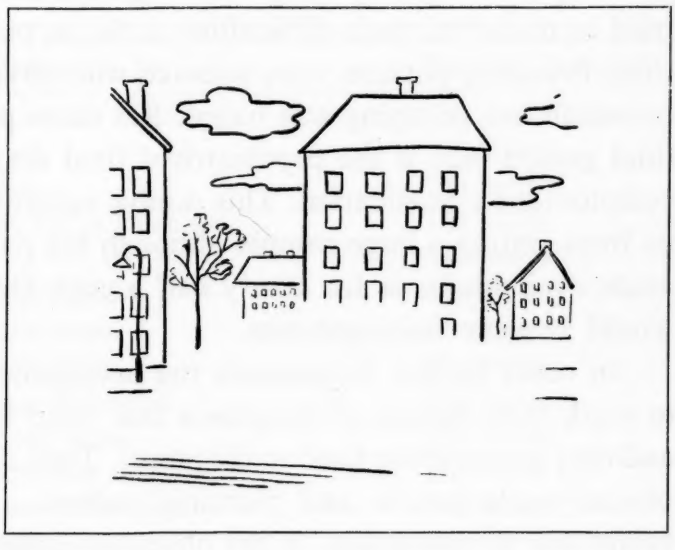
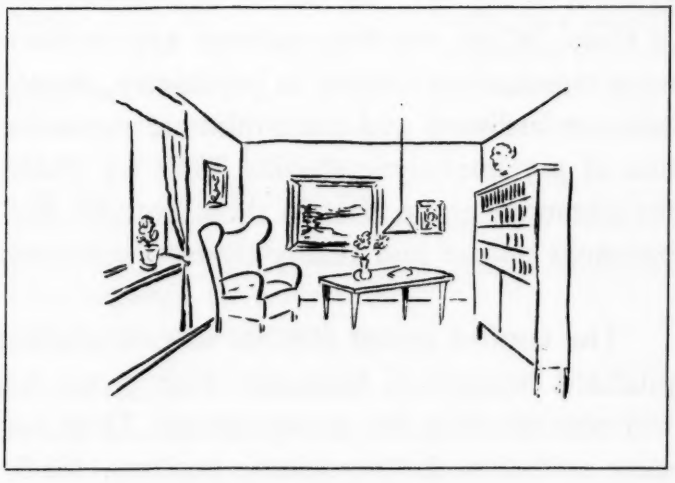


Fig. 3. The B-stimulus in the third series.



change was registered in that part of the series where B appeared alone five times at .05 seconds. Furthermore, it proved easy to decide when A appeared as an independent percept.

SUBJECTS

Any selection of groups of mentally disturbed Ss offers considerable problems as soon as psychiatric diagnoses are involved, since diagnoses differ not only between clinics but between individual psychiatrists as well. Even if the meanings of diagnostic concepts were well defined, diagnoses would nevertheless often have to be considered uncertain. A classification based on obvious symptomatic patterns would no doubt be more reliable. But the possibility of similar symptoms masking radically different etiologies would probably render purely symptomatic groups rather heterogeneous. We have tried to minimize these difficulties as far as possible by using double controls. Primarily patients were selected who obviously evinced the symptoms on which our grouping was based. But these patients were retained in the final groups only if the psychiatrists' final diagnoses also agreed with the symptomatic classification. This double screening of possible Ss prevented us from getting a large sample although the patients of two hospitals were made available to us for nearly half a year. But we hoped that the groups would become homogeneous.

In order further to promote the homogeneity of our groups we chose to work with classes of symptoms that were relatively well circumscribed and easy to recognize (and to diagnose). Thus our groups comprise hysteric, obsessive-compulsive and paranoid patients. All the hysterics displayed symptoms of conversion. In the obsessive-compulsive group no subject was as yet completely crippled by compulsive rituals, obsessive ideas, phobic systems or the like, but similar (anankastic) symptoms were obvious in all of them. When selecting patients with paranoid symptoms we did not, as most classical text-books in psychiatry, draw a sharp line of demarkation between ordinary and schizophrenic paranoia, thereby following the custom of psychoanalytic studies. Thus we added some psychotic patients to the paranoid group. But in these patients the delusions were mainly of a paranoid nature and reality-contact was otherwise relatively unimpaired.

The control group consists of orthopedic patients, most of them hospitalized because of fractures. This group may at least be considered as less neurotic than the groups above. Their social background is about the same as that of the psychiatric patients. All Ss are listed in Table I.

Table I: Subjects.

Group	< 30 yrs	30-55 yrs	> 55 yrs	sum
Paranoid	1	5	2	8
Compulsive	5	5		10
Hysteric	2	3		5
Normal	6	12		18
Sum	14	25	2	41

3 P, 4 C, 5 H, 5 N are women.

RESULTS

As an introduction to the presentation of the results, two protocols may serve to illustrate our basic data. Patient 1 is paranoid, patient 2 compulsive.

Exposure*	Patient 1 (series 1)	Patient 2 (series 3)
.050B	The same (i. e., the buildings).	The same (i. e., the interior).
.050B	No difference.	The same.
.050B	The same as far as I can see.	The same.
.050B	The same.	The same.
.050B	Just about the same.	The same.
.010A + B	Just about the same.	The same.
.0125A + B	Just about the same.	The same.
.016A + B	Perhaps some difference.	The same.
.020A + B	Very much the same.	The same.
.025A + B	The same also this time.	The same.
.032A + B	Could be somewhat bigger.	The same.
.040A + B	More roofs now.	The same.
.050A + B	It's bigger: the central building, the windows.	The same.
.064A + B	Still bigger now.	There was a car also, otherwise the same.
.080A + B	Central building very big.	Same as last time.
.100A + B	Parked cars. The house – wall behind.	Same as last time.
.125A + B	Seven parked cars. Otherwise the same.	
.160A + B	The parking lot wider, cars closer to me.	
.320A + B	Three big cars and the buildings.	
.640A + B	One car and the buildings. (This report is not changed at the two following exposures).	

* .050B = B presented for .050 seconds; .010A + B = A presented for .010 seconds immediately followed by B presented for .050 seconds, etc.

Table II: Exposures at which various phases of A appear.

Group	First appearance of an independent A		First correct A		Final correct A + B	
	Series 1	Series 3	Series 1	Series 3	Series 1	Series 3
Paranoid ..	5.8	7.1	7.6	9.0	9.3	10.1
Compulsive	4.8	6.1	5.0	7.2	5.5	7.6
Hysteric ...	5.2	5.2	5.2	5.6	6.0	7.5
Normal ...	5.3	6.1	5.7	7.8	6.8	8.0

In the protocol of patient 2, the perception of B does not change during the series. A appears as an end-stage (when exposed for .064 seconds) within B and remains the same. Patient 1 is quite different. The changes implied at exposures .016A and .032A cannot very well be distinguished from changes occurring in the control series (phases being "about the same"). At the next exposure, only stimulus details (roofs not previously reported) are added, but later (.050A, .064A, .080A) real change becomes evident. When A emerges at .100A the car is multiplied. At .160A B has changed again as well as A. Not until .640A and the following exposures are correct reports found. When the first car appears in the incongruent series 3 of this patient it is reported "in some sort of locality," including a summer-house, the car standing among bushes in front of it. The continuity of the perception of B, previously always related to the interior of a room in spite of obvious change, has now been broken as a result of the emergence of A.

Let us give a general orientation about the results by presenting the average exposure number (in the main series) at which various phases of A emerge (see table II). The first independent perception of A is most often registered at the fifth exposure in the congruent series 1, somewhat later in the incongruent series 3. A similar difference between the series also is obvious as regards the appearance of a first correct A or the final correct A + B (the first correct A + B in a sequence of three). Because of the considerable intra-group variation as regards the emergence of various phases of A we obtain no significant differences between the groups. The paranoid Ss seem to discover A a little later than other groups; but their high average values are partly due to the eldest paranoid S alone.[4] There is not only displacement between the first appearance of A and the first correct A, but also between the later phase and the final correct A + B. The compulsive and the paranoid Ss seem to differ in these respects, the displacements being greater among the latter.

4. In no other instance does age contribute to the group differences.

Table III: The number of exposures covered by the A-genesis.

Group	Series 1	Series 3	Series 3-1
Paranoid	4.5	4.0	—.5
Compulsive	1.7	2.5	+.8
Hysteric	1.8	3.2	+1.4
Normal	2.6	2.9	+.3

Comparisons: P and C in series 1, $P < 0.5$
N and C in series 1, $P < 0.5$

Because of the inter-group differences in displacement between various phases evident in table II, we have computed the duration of the A-genesis from the first appearance of an independent percept to the final A + B (see table III). The compulsives differed significantly from the paranoid Ss as well as from the normals in series 1, the differences being tested by the non-parametric test developed by White (13). Differences diminished in series 3. Preliminary inspection reveals that compulsives often inserted a number of uninterpreted representatives of A in series 3, whereas in hysterics, B undergoes a series of changes before a final correct A + B is established. These trends will be the object of a more detailed analysis below.

The first independent representation of A may be an end-stage, a correct A, or a pre-stage eventually leading up to an end-stage. Among the pre-stages we differentiate between uninterpreted reports ("a muddle of new lines," "something in the foreground," etc.) and incorrect interpreta-

Table IV: First appearance of an independent A-phase.

Group	Series 1			Series 3			Both series	
	Not interpreted	Correct	Incorrect interpret.	Not interpreted	Correct	Incorrect interpret.	No incorrect interpretation	Incorrect int.
Paranoid ...	2	4	2	2	3	3	3	5
Compulsive .	1	9	0	6	4	0	10	0
Hysteric	0	5	0	1	4	0	5	0
Normal	5	12	1	6	5	7	10	8
All	8	30	3	15	16	10	28	13

Comparison (corrected for continuity):

End-stage/pre-stage reports;
Series 1 versus Series 3 $\chi^2 = 8.36$ $P < 0.1$

Incorrect interpretations (one or both series)/other reports;
P versus C $\chi^2 = 5.82$ $P < 0.2$
N versus C $\chi^2 = 4.24$ $P < 0.5$
P + N versus C + H $\chi^2 = 8.80$ $P < 0.1$

Table V: The occurence of incorrect interpretations of A.

Group	Number of Ss with incorrect interpretation		
	Series 1	Series 3	One or both series
Paranoid ..... n = 8	3	5	7
Compulsive .... n = 10	0	1	1
Hysteric ..... n = 5	0	0	0
Normal ..... n = 18	4	8	11

Comparisons (corrected for continuity):
Incorrect interpretations (one or both series)/other reports;

P versus C $\chi^2 = 7.90$ $P < .01$
N versus C $\chi^2 = 4.93$ $P < .05$
N + P versus C + H $\chi^2 = 12.6$ $P < .001$

tions. Table IV summarises the initial reports on A. First of all, there were more correct reports in the congruent series 1 than in the incongruent series 3, a result in agreement with those mentioned in the introduction. But the groups also differed among themselves. A was more often represented by interpreted pre-stages in paranoid Ss and normals than in hysterics and compulsive patients, who either saw the end-stage immediately or started out with one or more uninterpreted pre-stages (compulsive Ss in series 3) A number of significant differences are given below the table.[5]

Those Ss whose first independent stage of A was left uninterpreted (or was an end-stage) sometimes developed (regressed into) incorrect interpretations later. The occurrence of interpreted pre-stages is summarized in table V. As before, normal and paranoid Ss differed from compulsives and hysterics in the number of Ss with interpreted pre-stages, such as "rugs" or "animals" (instead of a car) on the floor in the room in series 3. A few answers counted into the incorrect category imply that instead of seeing *one* car, S saw two or three, in one case up to seven. This concerns 4 Ss, one of whom with another incorrect report. Such a multiplication of A might follow a correct perception of it, as if the S, for instance, had to fill up the parking-lot created in front of the buildings (patient 1 above). Re-

5. The χ^2-values have been corrected for continuity. If the expected cell frequences were less than 5 or the distributions unsymmetrical, we controlled the value by using Fischer's exact 2×2 calculation. None of the P-values was found to be too low.

Table VI: The genesis of A.

Group	Continuous	Discontinuous
Paranoid	0	8
Compulsive ...	9	1
Hysteric	4	1
Normal	6	12

Comparison (corrected for continuity):
Continous/discontinous, P + N versus H + C
$\chi^2 = 13.0$ $P < .001$

ports that the car was moving (seldom given in series 3) did not affect the classification of A as an end-stage, however.

One way of summarizing the above computations would be to classify A according to its genesis. The categories of continuity and discontinuity used here have solely phenomenological connotations, and do not impeach the assumption of a continuity of motif (4) in each series. The continuous group thus includes those geneses where A either appears immediately as an end-stage or is preceded only by uninterpreted pre-stages. In the discontinuous category, a series of uninterpreted stages may be broken by stages where A is not reported at all or by interpreted stages; those series where an interpreted or correct A is followed by uninterpreted or blank stages are likewise included in the discontinuous group. As could have been predicted beforehand, compulsive and hysteric Ss differed from the rest of the sample in this respect (see table VI). Very little happened in the genesis of compulsives and hysterics as far as A was concerned, whereas all paranoid Ss and many normals displayed a variety of reactions. However, four out of five hysterics would be classified as discontinuous if changes in B could be regarded as a first representation of A, while all compulsives except two would still belong among the continuous Ss (see table VII).

When scoring change of B in the main series, we followed the procedure outlined in the introduction. Thus we registered only such changes of B as had no counterpart in the preceding control series; the reconstruction of independent stages of A was naturally not considered as a change in B. In some Ss the original B suddenly disappeared, leaving a blank stage or a blur of lines that was not interpreted. A completely new B or a new kind of B might also appear instead of the correct one. More often, however, the connection with the preceding B was preserved, only the distances within the picture, the size of objects or the contours being strikingly affected. Sometimes the S would report nothing but: "This is not at all the

Table VII: Change in B, one or both series.

Group	Change before A has appeared	Change only when A has appeared or after	No change
Paranoid	4	3	1
Compulsive	1	1	8
Hysteric	3	1	1
Normal	7	5	6

Comparison (corrected for continuity):

Change/no change; C versus P + H + N, $\chi^2 = 7.20$ $P < .01$.

old picture, but I cannot tell the difference." All these changes appeared either before the emergence of any independent perception of A, or at the same time as A appeared, or only after (see table VII). The difference between the compulsive group and the rest of the sample is significant with regard to change or no change in one or both series, most compulsives obviously refusing to see change in B throughout the entire experiment. But the category "change only when A appears or after" coincides with what has been termed cleavage in 2 paranoid Ss, 3 normal, and 1 compulsive. Repeating the comparison above, but excluding all Ss belonging to the middle column in the table, we obtain a χ^2 of 5.11 ($P < .05$). It is interesting to note that the one compulsive S displaying changes in B before the emergence of A was also characterized by very obvious projective mechanisms.

Assuming that changes in B before the emergence of an independent stage of A is produced by pre-stages of A beyond awareness, as was the case in the square experiment previously reported, we may likewise say that this change in B is the first representation of pre-stages of A. Then the reconstructed genesis of A should include all phases from the first change in B (or, the first independent A if there is no such change) until the final correct A + B. As shown by table VIII, which gives the average exposure numbers at which A was first represented, there are only slight differences between groups with regard to the beginning of the A-genesis. But differences as to the number of phases included in the genesis of A stand out even more clearly than in table III. The development of A was prolonged in paranoid Ss (who were also somewhat older) as against the compulsive where end-stages came to the fore immediately or soon after the reconstruction had started. It is also evident that changes in B very rarely occured in the very beginning of the main series, although B was likely to

Table VIII: The very first reconstruction of A (average exposure number for the final correct A + B within brackets).

Group	Series 1	Series 3
Paranoid	4.6 (9.3)	5.8 (10.1)
Compulsive .	4.7 (5.5)	6.1 (7.6)
Hysteric	4.6 (6.0)	4.2 (7.5)
Normal	5.2 (6.8)	4.9 (8.0)

be more unstable then than after a number of additional exposures. This helps to support the assumption that B-change is an important category of scoring in our experiments.

The mechanism of cleavage was mentioned at the beginning of this paper. The concept used here comprises all instances when a correct A does not emerge until A has in some way or other been separated from the original B. Many Ss accomplished this by separating A and B as to time of appearance, seeing A before (sometimes after) B on the screen. Neither was it uncommon for a subject to state explicitly that A appeared at the same time as B, but did "not belong to it at all." B might also disappear completely, or Ss refrain from interpreting it. Finally, some Ss reported a completely new B with no reference whatever to the foregoing one. In some Ss one mechanism of cleavage succeeded the other; but table IX is arranged on the basis of the initial mechanism. It is important to note that cleavage was used by 15 Ss in the incongruent series 3 as against only 4 Ss in series 1; such a result supports our interpretation of cleavage as a prerequisite to the acceptance of a correct A. The mechanism of cleavage was

Table IX: Cleavage (one or both series).

Group	A seen before B or not belonging to it	B disappears or is not interpreted	New B appears	No cleavage
Paranoid	4	1	2	1
Compulsive ...	0	2	0	8
Hysteric	1	0	0	4
Normal	2	2	1	13
All	7	5	3	26

Comparison (corrected for continuity):

Cleavage/no cleavage; P versus C + H + N, $\chi^2 = 8.55$ $P < .01$
Column 1/other columns, P versus C + H + N $\chi^2 = 5.00$ $P < .05$

used more often by paranoid Ss than by the other individuals. When we confine calculations to the kind of cleavage where the integrity of both A and B is preserved (the left column in table IX) we still obtain a difference between the paranoid and other groups. If it is argued that temporal cleavage is a consequence only of a late appearance of the correct A, we may rejoin that the most convenient mechanism of cleavage is always used. On the other hand, there is very little difference between initial temporal cleavage and other forms of cleavage with respect to the average exposure number where these mechanisms were first used (8.1 as against 7.6 in series 3).

DISCUSSION

The perceptual processes induced by this experimental technique will be handled only by means of such concepts as can be defined by the experimental operation.[6] Thus we distinguish between C-phases, i. e., the fairly correct and stabilized end result of a perceptual process (defined according to selected criteria), and P-phases leading up to (preceding) such a final phase (see further 4). Since the main experiment did not begin until S apprehended B correctly, the perception of B represents the adapted C-phase level on the basis of which pre-stages of A were reconstructed. This mechanism was particularly studied in the pilot experiments reported in the introduction. But the reconstructed pre-stages of A were more or less closely incorporated into (restricted by) the pattern of B, i. e., one important aspect of our analysis concerned the "independence" of the genesis of A in its early phases. The reality-context represented by B, on the other hand, was not always insensitive to the construction process of A. We have seen how it often changed as a result of pre-stages of A beyond awareness or disappeared when the A-process emerged for the first time as an independent percept, etc.

Compulsive individuals "cannot let themselves be surprised either by feelings or by perceptions that have not as yet been put into categories" (2, p. 288). They force the world into a definite system which should not be allowed to change. This defense-serving *quasi*-objectivity of theirs was probably reflected in our experiments when they continuously reported the original (objective) B throughout a series. The reconstruction of various

6. This is rarely the case in traditional S-R forms of theorizing within the field of perception where hypostatized variables (set, attitude, personality trait, etc.) are supposed to intervene between stimulation (which is often, and erroneously, defined as the average behavior of a sample of Ss) and perception (response).

phases of A did not concern the accepted B. At the same time there were, consequently, few and meagre P-phases of A, practically none of them interpreted into the context of B (series 3), or, the compulsives avoided categorizing A before they had reached definite certainty about its meaning. Similar results have been reported by Kragh (4), whose compulsive Ss avoided or *isolated* the unstable P-phases (cf. also the cumulatives of Smith and Klein, 11). Since the integrity of B seemed to be guaranteed by these mechanisms (of isolation), we were not surprised to find that few compulsive Ss used the mechanism of cleavage (B disappears when A emerges, is seen before or after A, etc.), even provided that this cleavage often served to reduce the conflict between an accepted C-phase and a new (incongruent) C-phase emerging.[7]

The majority of other Ss tended to differ from the compulsive group as far as the influence on B by P-phases of A were concerned. In the paranoid group reality-relations are generally frail, as the defensive use of projected delusions shows. The world of many hysterics is inclined to change perpetually. We therefore expected (drastic) change in B. But even among normal individuals B often changed as a result of the reconstruction of pre-stages of A (cf. 10 and the introduction), probably because the very security of reality relations makes change possible – even if not as pronounced as in other groups. At the same time normals were not inclined to separate an independent P-phase of A (or a suddenly appearing C-phase) from B as often as paranoid Ss, who may have to guard themselves against perceptual conflicts by "splitting" A and B, by accepting only A (perhaps in a "delusional" B), etc. In hysterics such a cleavage did not seem to be necessary either; let us presume tentatively that total avoidance of A is a substitute defense possibility for them. As far as incorrect interpretations of A are concerned, we expected and registered them in paranoid patients and normals, whereas in hysterics pre-stages of A appeared more often to be inhibited.

The change of B should be considered as one of our most important findings, not least the change registered before the appearance of an independent P-phase of A. This phenomenon could be described as a compromise between B and the beginning A-process, the continuity of B being preserved at the same time. Changes in B were more pronounced and covered more phases in people with unstable reality-contact, i. e., paranoid

7. Incidentally, in the (academic) pilot group, more Ss displayed these compulsive symptoms than in the (normal) control group. Since compulsive Ss disqualify P-phases and report only C-phases, the exclusive use of academic Ss would tend to support traditional S-R theories (cf. 8, 9).

Ss and some hysteric, than in normals where they seemed to be of a more occasional character. The influence on B was rarely scored when the first independent P-phases emerged, however, which were instead often compromised toward B in Ss where B-changes occured. When the uncompromising C-phase of A came to the fore (late as compared with the very first representation of A in paranoid individuals) B changed again in many Ss. The new B was most often directly related to the old one. But in a number of Ss, especially paranoid, it disappeared or was completely altered. The latter reaction was classified as cleavage because A did not present itself within the original B. While other forms of cleavage implied that the C-phase of A could be accepted and the integrity of B preserved at the same time, the appearance of A as a C-phase in these instances caused a complete break in the continuity of B (regression). One might say that the new perception could be stabilized only because the old one was sacrificed. But a conflict between the original B and an incongruent A on the verge of emerging was at least averted for the moment; and the original B could easily reappear later in the series (often temporally separated from A).

The findings presented so far do not exhaust the possibilities of our tachistoscopic technique; they rather initiate a long series of further experiments (cf. 3). As a matter of convenience, our pictures have been simple drawings and our classifications mainly of a formal order. It would otherwise have been impossible to ascertain the significance of phenomena within a completely new field of investigation. However, the future application of our technique should be changed in two respects in order to ensure response categories suitable for content analysis and relevant for central adaptive and defensive mechanisms. Thus pictures including human beings will be substituted for the present stimuli, and A will be manipulated in such a way as to prolong its genesis. Then it should be possible to carry out the phase-to-phase analysis implied by the actual-genetic model.

SUMMARY

Two stimuli, A and B, were presented immediately after each other in a tachistoscope, B at the threshold value for correct perception, A at a subthreshold value subsequently raised in geometrical progression. In one series A might be congruent with B (car + street-scene), in the other series the two stimuli were incongruent (car + interior). A pilot study demonstrated that the first independent representation of A was affected by the perceived context of B, i. e., the reconstruction of initial pre-stages of the correct percept would be partly organized on the basis of the percept already established. But pre-stages of A beyond awareness might also exert a significant influence upon the perception of B. The genesis of stimulus A from its first appearance

either as an independent percept, or as an obvious influence upon B, seemed to reveal consistent patterns reflecting mechanisms of adaptive regulation.

Part of these experiments were applied to groups of normal (n = 18), compulsive (n = 10), paranoid (n = 8), and conversion-hysteric (n = 5) patients. The results of the pilot study guided the scoring. Compulsive Ss held on to an unchanged B throughout a series. The genesis of A was often contracted to a correct end-stage, uninterpreted pre-stages sometimes preceding. In the other groups B often changed, either before A had appeared as an independent percept, at the same time, or after. Change was most obvious in paranoid Ss. In contrast to both compulsive and hysteric Ss, the genesis of A in normal and paranoid Ss often included interpreted pre-stages. Sometimes S did not accept the correct A until it had been separated from B in one way or other. Paranoid Ss, especially, employed this mechanism of cleavage; it might even happen that the correct A emerged in a "delusional" B.

REFERENCES

1. Bruner, J. S., & Postman, L. On the perception of incongruity: a paradigm. *J. Pers.*, 1949, *18*, 206-23.
2. Fenichel, O., *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
3. Klein, G. S., Spence, D. P., Holt, R. R., & Gourevitch, S. Preconscious influences upon conscious cognitive behavior. *Amer. Psychologist*, 1955, *10*, 380.
4. Kragh, U. *The actual-genetic model of perception-personality*. Lund: Gleerup, 1955.
5. Sander, F. Experimentelle Ergebnisse der Gestaltpsychologie. *Ber. 10. Kongr. exp. Psychol.* (in Bonn 1927), 1928, 23-88.
6. Smith, G. Development as a psychological reference system. *Psychol. Rev.*, 1952, *59*, 363-69.
7. – Twin differences with reference to the Müller-Lyer illusion. A study in modal summaries and serial change. *Lunds Universitets Årsskrift, N. F.*, Avd. 1, 1953, *50*, 3.
8. – The place of physiological constructs in a genetic explanatory system. *Psychol. Rev.*, 1954, *61*, 73-76.
9. – Freud's theory of perception. (To be published).
10. Smith, G., & Henriksson, Maj. The effect on an established percept of a perceptual process beyond awareness. *Acta psychol.* 1955, *11*, 346-355.
11. Smith, G., & Klein, G. S. Cognitive controls in serial behavior patterns. *J. Pers.*, 1953, *22*, 188-213.
12. Smith, G., & Kragh, U. Do micro-genetic sequences reflect life history? An illustration. *Acta psychol.*, 1955, *11*, 504-12.
13. White, C. The use of ranks in a test of significance for comparing two treatments. *Biometrics*, 1952, *8*, 33-41.

THE NATURE OF G AND THE CONCEPT OF INTELLIGENCE

The Idea of Complementarity in Psychology

ALEXANDER W. STERN

Institute for Theoretical Physics, University of Copenhagen

The general features of nature lend themselves to expression in symbolic and mathematical form, and, by deducing the implications of their symbolic pattern, one is enabled to pierce into the unknown. Often, general considerations of this kind lead one to the planning of decisive experiments and observation, and our knowledge and understanding of nature is thereby increased. Here, in a somewhat introductory fashion, a general theory of the intellective factors of the mind is sketched. Although speculative, it provides a framework for a basic interpretation of the fundamental concepts underlying intelligence testing and that branch of psychology concerned with an understanding of the factors of the mind. Also, this paper has been written to stimulate further inquiry, experimentation, and observation in a field of theoretical and practical importance.

Although the primary intellective factors[1] may be consistently identified, their behavior is not exclusively of an independent nature. Moreover, they cannot be consistently treated by any mathematical-statistical scheme as if they existed as purely isolated factors. The implications of analyses of intelligence tests are in accord with the above statements.

I assume that mental phenomena, like other natural phenomena, may be analyzed into basic conceptual elements which may be designated as the primary intellective factors. They, in themselves, do not possess the properties of mental abilities, but, rather, the potentialities for mental behavior which give rise to the actualities of mental abilities upon suitable environmental stimuli.

We see why process – environmental interaction – is of such basic importance in nature, for it converts possibilities or potentialities of behavior which reside in the isolated units of nature, such as, the elementary particles of physics, the genes of biology, or the factors of psychology, into actualities. The important and relevant physical, biological, or psychological properties are not pre-existent in these units.

The possibility arises that the factor of mental phenomena may exhibit a *collective* as well as an *independent* behavior. The two modes of behavior

1. What these primary intellective factors are is not pertinent to this discussion, and I accept the consensus of psychologists as regards their identification.

form a fundamental *complementarity suggestive* of the complementarity exhibited by the well-known wave-particle duality og quantum theory. The idea of complementary pairs of opposites seems to stem from some fundamental aspect of nature, for its meaning and sense can be generalized to apply to wide domains of knowledge. The principle of complementarity, generally speaking, refers to a manner of describing a domain of phenomena that exhibit opposite and contradictory modes of behavior, such as the particle and wave behavior of the elementary particles of nature.

This particle-wave duality arises because one cannot separate the behavior of the elementary particle from its interaction with the system representing the apparatus. On the atomic level, the apparatus is not the passive instrument or system of classical physics, but constitutes the experimental conditions under which the phenomena occur. The physical properties of the electron, for instance, should not be considered as intrinsic properties belonging to the electron. Rather, they arise out of its interaction with other systems such as the apparatus. Consequently, different experimental systems interacting with the electron will give contrasting, even contradictory, pictures of the electron's behaviour. This paradox was epistemologically resolved by Niels Bohr in his Principle of Complementarity aforementioned. It is a description which employs the use of contradictory concepts in a complementary relationship. The contradictory aspects of the phenomena cannot, of course, be simultaneously observed; rather, they are mutually exclusive. However, there is a glass of phenomena in the domain which demands, for its logical description and understanding, the employment of one of these concepts, and there is another class which demands the employment of the antipodal concept, so that a complete description of the entire domain of the phenomena necessitates the use of both concepts in a complementary manner.

Mental phenomena will exhibit contrasting properties under different environmental situations. Under those environmental situations which call for the exercise of singular, specialized intellectual abilities, there will occur an independent-like behavior of the mental factors involved. Under more general situations, such as those that one meets with in a social environment, or in meeting the practical problems of living, politics, or leadership, where, if solutions do exist, they are not unique, there will occur a collective-like behavior of the mental factors. Thus, the environmental situation of the mind plays a role similar to that of the experimental situation in physics under which the phenomena occur.

Factors are not found existing by themselves. As such they are but conceptual idealizations. They are always found associated with other

factors, but this association does not destroy their independent character. Consequently, they can be isolated and identified. In intellective skills or abilities, certain single factors obviously predominate, and the other factors associated with the predominant ones may be considered in the manner of secondary factors. The independent-like behavior of the factors is manifest in the exercise of the various mental abilities and in learning.

The collective type of behavior suggests a coupling of the various intellective factors, which gives rise to resonance modes of operation of these factors. That is to say that there is a type of reinforcement of the intellective elements or factors when acting in a collective manner, that gives rise to a new pattern, a new totality of behavior that can be identified with the manifestation of intelligence. The peak of this resonance is a measure of the degree of intelligence. To state it differently, intelligence, in this conceptual framework, is the resultant collective behavior arising from the various couplings among the intellective factors. The strength of this resonance or collective coupling I shall designate as g. This resonance or collective coupling is not to be confused with the fact that factors are always found associated with other factors. This association does not destroy the independence of the factors and has nothing to do with intelligence. Intellective abilities can be shown to contain several independent factors. Intelligence, on the other hand, cannot be unambiguously decomposed into separate elements or factors. The absence of g coupling or resonance among the intellective factors is the condition for their independent behavior. The stronger the g coupling, the better defined is the collective behavior of the intellective factors and the better the definition of the intelligence pattern. The existence of the g coupling thus limits the assumption of independent factors of mental behavior and signifies the complementary aspect of their collective behavior.

Intelligence, as here defined, possesses gestalt, or configurational properties which are of considerable importance. The gestalt, or configurational, properties give rise to an intelligence pattern. The features of this pattern, the manner in which the intellective factors are coupled, that is, the symmetry or uniformity in the coupling strengths, are as important for intelligent behavior, for social effectiveness, and for living as is the quantitative aspect of intelligence, which is concerned with the degree of the intellective factors that go to make up intelligence. The strength of the g couplings and the uniformity in their strength, which are basic in determining the symmetry properties of the intelligence pattern, are conditioned by traits other than intellective, such as, interest, drive, incentive, perseverance, and, most important, those personality factors which give rise to

what may be called adult maturity. Intelligence is thus a fundamentally different concept from intellect. What is relevant regarding intellect is essentially quantitative. Intellectual ability is a matter of "how much", while with intelligence it is relevant to ask for qualitative features as well. The kind of intelligence one possesses may indeed be as important outside of the classroom and in practical situations as the amount of intelligence. In a personal communication to the writer Dr. David Wechsler stated that Binet I. Q. tests cannot be used as the sole criterion for deciding whether a subject should or should not be committed to an institution for mental defectives. He states that "one child with an I. Q. of 75 may be definitely defective, while another with an identical I. Q. or, indeed, one with 5 or 10 point less may be far from so classifiable." Dr. Wechsler pleads for a broader concept of intelligence to include non-intellective factors that make for social adaptiveness. It is the writer's contention that the variables responsible for social adaptiveness – the non-intellective factors – influence the intelligent pattern. The more asymmetrical the pattern, the more ill-adjusted is the subject. Of course, there must be a reasonable minimum of intellective factors to constitute an observable intelligence pattern.

Summarily, one may note that, although g is involved in all manifestations of human intelligence, it is not some identifiable thing, some common or super-intellective factor. It is the coupling responsible for the collective behavior of the constituent elements or factors of mental phenomena, that gives rise to a new level of mental organization. It is on this level of organization that the concept of intelligence emerges.

It was noted above that the strengths of the g couplings and their uniformity in strength may be conditioned by traits other than intellective, and mention was made of those personality factors which give rise to what may be called adult maturity. It is common observation that the development of this type of maturity with age is radically different from intelligence as measured by I. Q. tests. It is in evidence much later, it exhibits a much slower growth, and levels off at a later age before falling. Moreover, it does not seem to be as fixed in its general development. It depends, in good part, on the interplay of environmental and hereditary factors. Although non-intellective factors may play an important role in intelligence, they play little or no role in the exercise of pure intellect.

Intellect, of course, is a function of the separate behavior of the primary intellective factors which are essentially of a hereditary nature. The curve of intelligence, as here defined, will consequently rise and fall much more slowly with age than do the separate primary intellective factors or the learning process which is a function of the independent behavior of

these factors. The importance that the above has in the study of geriatrics is, of course, obvious.

It would be of interest and importance to devise tests that actually measure intelligence. Of course, it is not advocated that such tests should replace the intelligence tests now in use, that measure essentially learning ability. These tests are most valuable for schoolroom use. Besides, the intellective and non-intellective abilities as well are essential for the acquisition of particular skills and techniques basic to the achievement of competence or excellence in the remunerative as well as the learned professions. However, even the layman can discern that excellence in any particular line of endeavor, since it requires a specialized singular ability, does not necessarily imply a concomitant high degree of intelligence.

What is emphasized here is the dual and complementary aspect of the intellective factors of the mind. This dual and complementarity aspect is exhibited as the independent-like behavior of the intellective factors as separate abilities and as their collective-like behavior as intelligence. There must be a sharp logical delineation between these two aspects of mental phenomena. For instance, as mentioned before, interest may condition the strength of couplings between the intellective factors, therefore, it is allied to their collective behavior. On the other hand, competence in a field of endeavor is determined essentially by the independent behavior of the intellective factors. Therefore, it is not correct to measure a subject's fitness for a profession by his interest in that profession. He may have an intelligent interest in a profession and yet lack the necessary degree of those intellective abilities essential for competence in that profession. The separate, or independent, behavior of intellective factors gives rise to an intellective ability. It is for this reason that aptitude tests which confuse these two complementary aspects of mental phenomena are not worth very much.

Consequently, one should not equate the marked possession of a mental ability to learn or excel in some specialized subject with intelligence. There is indeed a correlation between the two, but not an equation. The intellect is paramount for the solution of specialized problems, but intelligence is essential for coping with the problems of living. When living takes the place of learning, then adaptive potentialities are as important as learning potentialities; social effectiveness as important as excellence in the acquisition of some intellective or non-intellective skill or technique; and intelligence at least as important as intellect.

For testing and measuring such potentialities, one should not employ tests that measure intellective abilities. Such tests are of little or no value

in the measurement of intelligence. Rather one should employ those tests that are designed to ascertain the extent to which the intellective factors are coupled to form a well-balanced integration. For instance, verbal tests that measure verbal understanding rather than a facility with words and number tests that exhibit the subject's ability to reason quantitatively rather than his facility with numbers should be supplemented by a mixed test that would test the subject's ability to give verbal expression to his reasoning processes and his insight into quantitative relations. Such mixed tests could yield an estimate of a subject's intelligence with words and numbers, and could be of value when combined with comprehensive tests, as cited above. A high score on some intellective ability or skill and a low score on both a comprehensive test and a mixed test would indicate that the subject's mentality is dominated and colored by his possession of that skill. It would suggest that his thinking is, in reality, a triumph of technique over intelligence. While the exercise of a fully developed skill or technique, be it of an intellective or non-intellective nature, may not be of a fully conscious nature, this is not so with intelligence. Intelligence involves awareness. There is no subconscious activity associated with the use of intelligence. The whole man is here in control of his mental faculties.

It has been known for some time that it is not possible to account for the total amount of variance in intelligence tests in terms of a reasonable number of factors. Wechsler has stated, "The greater the number of different tests included, the smaller generally the total percent of variance accounted for, and this is seemingly independent of the number of factors extracted." In fact, he states that factors beyond the first 3 or 4 usually contribute so little to the already accounted for variance, that it is generally not profitable to extract them. He claims also that the amount of variance accounted for is seldom more than 60 percent of the total.

It is the writer's contention that no amount of additional factors of any nature whatsoever, intellective or non-intellective, can account for the total variance. The variance cannot be ascribed to any kind of independent factor or factors. What has been overlooked in intelligence tests is not any factor or factors, but the existence of intelligence itself. It is remembered that intelligence is a manifestation of the collective behavior and integration of the intellective factors of the mind, and that the coupling between the factors may indeed be affected by non-intellective factors, such as interest, drive, temperament, and the like. It is satisfying to know that intelligence tests do not, or, rather, cannot disregard intelligence altogether.

In order to achieve a consistent formulation of the factors of the mind, it may be of prime importance to distinguish unambiguously between those

non-intellective factors that effect the integration of the various intellective factors and are, consequently, a part of the intelligence pattern and those non-intellective factors that may be termed physiological factors, be they normal or abnormal, and that are important in determining the effectiveness of intelligent behavior. The first type may be called the non-intellective psychological factors, and consist of such elements as interests and those aspects of perseverance and temperament that are related to mental achievement. The second type consists of basic anxiety, undue nervousness, emotional imbalances, and the like. It is obvious that this type may also influence I. Q. scores to various degrees.

The two types of non-intellective elements spring from different causes, one being a matter of the cultural environment and the other a matter of one's inner environment – of body chemistry. Consequently, the first type may be conditioned by education, while the other type may spring from organic causes and demand the attention or the co-operation of the psychiatrist. The non-intellective psychological factors have a positive influence on intelligence, while the physiological factors have essentially a negative influence, that is, they hinder achievement.

Of course, one may say with complete justification, that the sciences, mathematics, physics, chemistry, as well as biology, physiology, and psychology, are all interrelated. Chemistry, for instance, may, in the not-too-distant future, play a basic role in the treatment of mental disease; and mathematics may play an important role in biology – in the basic problem of replication and in the structure of proteins. However, in order to avoid utter confusion, it is imperative that each science retain its integrity. Only when sciences are in an advanced stage of evolution and development may they be profitably combined.

The paradox caused by the learning curves which express the growth of the various intellective factors with age, if taken seriously, would lead one to equate the intelligence of the average adult to that of a teen-ager. Obviously, those curves are not at all relevant for testing people who have already demonstrated their learning ability or lack of it many years back, and who have acquired some professional skill or technique. What is relevant to know about adults is the degree of integration and balance of their mental abilities. The writer believes that this symmetry, or balance, of their mental pattern yields a better measure of those traits which are of importance in the adult group. It is his opinion that there exists a strong correlation between the balance displayed by one's personality pattern and the symmetry of one's mental pattern. The social effectiveness of one's intellectual abilities is conditioned by their integration into a symmetrical

functional unit. It is the phenomenon of this integration that should be identified with intelligence. The integration of the primary intellective factors is a much slower process than their separate development, and is not directly influenced by hereditary factors.

It is, indeed, an all too human error to identify something that we believe to exist with some thing – with substance rather than form. In fact the very structure of our language predisposes us to make such an identity, and channels our thought in this direction. However, contemporary science reveals that reality is dynamic rather than static. Form gives rise to pattern, and from pattern emerges process. Although g cannot be identified with some higher intellective factor, with some entity, it is, nevertheless, just as real.

As was previously noted, factors of the mind are not absolute. They always exist in association with other factors. In the anthropomorphic terms of classical psychology, one may say that this fact is a reflection of the intimate association of meaning with thought. For example, if one wishes to exhibit the atomistic unit character of memory, one devises tests that resort to meaningless words and symbols. However, the exercise of memory in human thought where it no longer behaves as a separate faculty is associated with meaningful content. The very concept of thought points to the connectivity of the mental faculties which the factors isolate and identify.

The difference between the memory of a giant computing machine and human memory is not that machine memory is a simulated one. On the contrary, machine memory should exhibit a closer correspondence to the concept of memory than human memory. This is because concepts have a rigorous and unambiguous meaning only when applied to systems which may be validly regarded as static, deterministic, and closed systems, and the machine is such a system. In fact, when human memory functions in a manner resembling that of a memory machine, its behavior is manifestly artificial. Human memory has an "open" character, and, consequently, cannot be considered as behaving in a single, isolated manner, but rather as being associated with other intellective factors signifying a level of mental organization that can be identified with the phenomenon of mind.

It is a pleasure to express my gratitude to Prof. Niels Bohr for his kind hospitality and encouragement and to Dr. David Wechsler for many illuminating discussions in the past.

ON THE 'CORRECTION' AND 'NON-CORRECTION' METHODS OF LEARNING

J. M. von WRIGHT
University of Helsinki

I.

The general issue of the correction vs. the non-correction methods of learning was brought into the limelight by Tolman in 1937 (10). The following year Hull and Spence (3) discussed these methods on the basis of conditioning principles, predicting that the C (correction) method would be more favourable in discrimination learning situations with animal subjects. They also presented experimental data confirming this prediction. However, Seward (8) has obtained opposite results in a fairly similar situation, finding much faster learning under NC (non-correction) conditions. Later experiments relevant to the issue in question (e. g. 6, 13) have given rise to some controversy, but have failed to yield unequivocal results, and at present no general conclusions as to the relative advantage of these two methods seem possible even for simple single-choice situations.

Still less is known about these methods as applied to human learning, in spite of recent emphasis on the problem (e. g. 4). In view of the vagueness of most of the current discussion about the methods, an attempt is made here to propose some distinctions, and to suggest some conditions which may be of importance for the relative efficiency of these methods in *human serial learning*.

When applied to a serial situation, the term 'non-correction method' has been used in two different ways, referring to (a) a situation in which S has to start from the beginning each time he makes any error, or to (b) a situation in which S on each trial proceeds with the task from the beginning to the end without (overtly) correcting any errors. In simple discrimination learning tasks with animal subjects these two variations of the method usually coincide, but they differ when there is more than one choice or part-task per trial. We shall here refer to the *latter only* as the NC method (for a discussion of the former, see (14)).

A straightforward comparison of the C and the NC methods in serial choice situations is possible only with two-choice tasks, or – more generally – with tasks where 'in principle' as much information can be obtained from a wrong choice as from a correct one. If this is not the case, there being, say, three alternatives for each item only one of which is correct, and if S learns by the NC method and makes an error, there still remain

two possibilities, either of which may be the right one. With the C method, on the other hand, S is given unequivocal information as to the correct response for each item. Consequently, it seems likely that the relative efficiency of the latter method increases with the number of possibilities at each serial position. However, the actual relationship is not known, for any particular type of task or situation (some information of a general nature may be found in (9) and in other papers by Thorndike).

The problem is somewhat different in traditional serial verbal learning experiments using the serial anticipation method and the method of paired associates. The number of possible responses at each step is here great, and S is in each case presented with the correct alternative after he has made his response, independently of its correctness. Thus the same information is given to S in the same way irrespective of whether the C or the NC method is used, the only difference being in the S's response, in that with the former method S is requested to pronounce the correct alternative aloud if he has made an error, while this is not the case with the latter method. The difference between the methods seems to be relatively unimportant, and it has frequently been ignored. Buxton and Bakan (1) report that the C method was more favourable to the occurrence of reminiscence, but it is doubtful whether their results permits of such an interpretation. Their results are moreover inconsistent with those obtained by Ward (11), and the issue remains undecided.

A distinction between the C and the NC methods is of greater interest in the field of serial sensori-motor and motor learning. Considering for a moment possible predictions from current learning theories, it appears that neither Hull and Spence's (3), nor Seward's (8) analyses of the discrimination learning situation can be applied to serial learning in their present form, but a general application of S-R reinforcement principles would seem to lead us to expect the C method to be more economical (except, perhaps, in the time scores). However, a specific and unequivocal prediction for a serial two-choice situation can be made from Guthrie's theory. The prediction is that the C method should be superior, and more superior the more distinctive and different from each other the external and internal stimuli with which S is 'faced' at the moment of choice at different choice-points, and the more constant the stimulus situation at each choice point from trial to trial.

This prediction is based on the three fundamental postulates in Voeks's formalization of Guthrie's theory (as quoted in 7, p. 364):

(1) Whenever a stimulus is contiguous with a response, it becomes maximally associated with that response.

(2) Whenever a stimulus, previously associated with a given response, accompanies another response incompatible with the first, the previous association is completely eliminated.

(3) The probability of a particular response occurring at a given time is an increasing monotonic function of the proportion of all stimuli present that are conditioned or unconditioned cues for the response in question.

The argument runs – briefly – as follows. Let K be a choice point in a serial two-choice task, and *Sk* the set of stimuli characteristic of K which – at least partly – impinges upon the subject S each time he comes to K. Let us further assume that S comes to K for the first time, and that he makes an incorrect choice, say, to the left. The total stimulus situation S_1 'accompanying' S's left-side choice contains *Sk*. If S now is learning by the *NC method* he will not return to K – and consequently not be 'faced with' *Sk* – until the next trial. When he on this next trial comes to K, the total stimulus situation S_2 (containing *Sk)* will tend to be followed by a repetition of the previous choice, i. e. by a repeated error, because of the similarity of S_2 to S_1. If, on the other hand, S is learning by the *C method,* he returns after his initial wrong choice on the first trial to the choice point K, where the new total stimulus situation S_3 (containing *Sk*) now 'accompanies' a right-side choice. When S on the next trial comes to K anew, the total stimulus situation S_4 is similar to both S_1 and S_3. S should, however, tend to go to the right (correct) side, as this was the most recent movement 'accompanied' by *Sk* (cf. postulate 2). Thus the C method should be more favourable to learning than the NC method on units on which S had made an error on the previous trial, while there should be no difference on units at which a correct choice had been made on the previous trial – the net total being in favour of the C method.

It is clear that the total stimulus situation in each case contains much more than *Sk,* but this is the case irrespective of whether the C or the NC method is used, and there seems to be no reason *within the framework of the Guthrie-Voeks's theory* to postulate any systematic differences between the two methods in this respect.

In an experiment made by the writer (cf. 14) learning by the C vs. the NC method was compared using a 12-unit linear pattern as (a) a finger maze, and (b) a stylus maze. With the former maze no significant differences were obtained between the two groups of subjects (the scores of those using the NC method being consistently but only slightly better). With the stylus maze the NC method proved significantly better in trial and time scores, the difference in errors being of borderline significance. The results thus tended to go contrary to the prediction from the Guthrie-Voeks's theory discussed above. A further finding was that while the error distributions were very similar for both finger maze groups, they were radically different for the stylus maze groups. On the basis of an analysis of this experiment it is suggested that in *serial* two-choice situations at least the following three considerations are of importance for the relative economy of the two methods.

(A) When there are only two choices, a wrong one and a correct one, information showing that one of them is wrong is logically equivalent to

information showing that the other one is correct. However, whether S will learn the correct choice equally well in both cases is a function of a number of conditions, the most important of which seems to be the 'complexity' of the task. In recent studies of concept formation (2), there is evidence that Ss consistently learn much less from 'all negative' instances than from 'all positive' ones, even when the amount of information is equated for the two types of instances (cf. also 12). However, the simpler the task and the more clearly defined the alternatives, the smaller is the difference. Another possibly relevant condition is the way in which the information is given. In studies of 'positive' vs. 'negative guidance' it has usually been found that emphasizing the correct response is somewhat more effective than emphasizing the wrong one, even when the task is very simple and well-defined. A comparison of the outcome of various experiments seems to indicate that this difference is larger using the NC method than when the C method is used (cf. 14), but no definite conclusion is possible as the two methods were never compared in exactly the same set-up.

(B) The relative efficiency of the two methods is also a function of the difficulty of the responses involved. If the execution of the responses *per se* is facilitated by making them during the learning period, then the additional training provided by the C method should tend to make this method more favourable. If, on the other hand, the responses involved are very easy, or if S already masters them, there should be little advantage gained simply from their execution; the very fact that negative practice is often effective in skilled performance when the correct part-responses are well known (5, pp. 291-295) indicates that sheer repetition of a response under these circumstances hardly aids the learning of the serial performance.

On the basis of these two considerations one might suggest that the C method is more favourable the more 'complex' the situation, and the greater the difficulty of the responses to be performed. As a limiting case – when the alternative responses are well known or easy to perform, and when they are so clearly defined that the transition from a negative to the corresponding positive instance does not present any difficulty – we would not expect much difference between the two methods.

There is, however, yet another consideration to be taken into account. (C) In several serial tasks, particularly in highly skilled ones, learning may be facilitated if S is allowed to perform the task as an uninterrupted sequence – even if it contains errors – whereas a correction of each error tends to disrupt the performance, preventing the organization of the task 'as a whole'. Confining the discussion to simple tasks, the C method should

be relatively less favourable the less each item (correct response) has characteristics which make it possible to learn it separately from all other items, e. g., the more the items (responses) in the series are defined solely in terms of their interrelationships. In the extreme case, when the same two alternatives are presented throughout and the task only consist in learning the sequence of the correct alternatives (e. g., a linear maze), correction of the errors adds very little information, but tends instead to interfere with learning. Such disruption of the smooth flow of the performance has greater detrimental effect the more attention the corrective activity requires, the longer time it takes, and so on. It may be pointed out, however, that in the experiment cited above in which the NC method was on the whole found to be more favourable, a change to the C method at a late stage of learning when only a few very persistent errors 'remained' was found to accelerate learning by emphasizing, as it were, the points at which S had to change his fairly rigid behaviour pattern.

II.

Of the three considerations presented above, the third seems to be the most debatable one. At the same time it is of some practical interest, as the C method is predominantly used in teaching motor and sensori-motor skills. Because of this the following prediction made on the basis of it was experimentally tested: In learning a sensori-motor task of the paired associates type, the NC method is expected to be as good as, or more economical than, the C method when the serial order remains *constant* throughout learning, while the C method is expected to prove decidedly more efficient when the serial order is *varied* from trial to trial. (In the latter case only the two first considerations, A and B, would apply.)

Apparatus and procedure. The apparatus, a simple temporal maze, is shown in outline in Fig. 1. It consists of a board into which a groove 7 mm in width is cut.

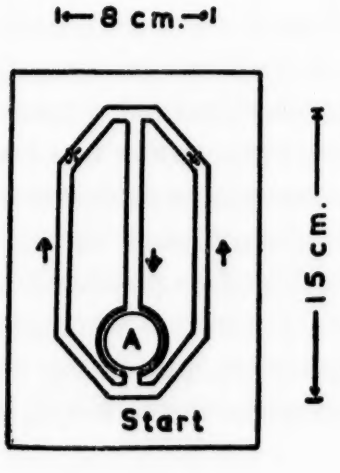


Fig. 1.

S followed the groove with a brass stylus, choosing at the starting point either the right or the left side alley, returning in each case along the middle alley to the starting point. In both side-alleys a piece of brass was sunk into the bottom of the groove (X' in Fig. 1). When the stylus touched either strip a circuit was closed. This circuit was used for recording the choice, and also, when S had chosen the wrong alley, a buzzer connected to the circuit gave a loud signal.

A round hole, 3 cm in diameter ('A' in Fig. 1), was drilled through the board close to the starting point, and covered with a sheet of glass level with the bottom of the groove. A shutter was so arranged that it covered this 'window' as long as S was traversing the maze, but leaving it visible when S was at the starting point or close to it. Twelve simple figures (15 mm high) were drawn on cardboard and shown one at a time in the 'window', the figures being changed by a simple device while the 'window' was covered by the shutter. The figures were all simple and chosen so as to give no directional cues of left or right; a sample is presented below:

S sat comfortably in a chair in front of the table on which the maze was placed, tracing the maze with his preferred hand and at a free choice speed. He was told that a number of figures would be shown in the 'window' one at a time. Each such figure was associated with either the left or the right side alley, and S had to learn by trial and error which side went with which figure. At the end of each trial a small lamp in front of S lighted. Ten seconds later the light went out again and the next trial began. The learning criterion was two successive errorless trials.

Four experimental conditions were used:

A : NC method, serial order constant
B : C method, „ „ „
C : NC method, serial order variable
D : C method, „ „ „

When the C method was used, S returned to the starting point each time he made an error, as soon as he heard the buzzer signal. Having been presented with the same figure a second time, he continued along the correct side-alley, returning to the starting point to make his next choice along the middle alley. With the NC method S always continued straight along the alley chosen, irrespective of its correctness, and returned to the starting point along the middle alley.

Under conditions *A* and *B* (serial order constant) the figures were presented in the same order on each trial for each S, *this fact being stressed in the instructions.* The correct serial order of the choices was LRRLRLLLRRL. – Under conditions *C* and *D* the order of presentation of the figures was randomly altered from trial to trial (systematic randomization being used). This, too, was stressed in the instructions. The same twelve figures were used in each case, each of them appearing once during each trial.

48 students in two elementary psychology classes (19-24 years of age) served as Ss. None of them knew the purpose of the experiment, nor had they participated in more than one learning experiment (see below) prior to the present one. They were divided into four matched groups on the basis of their scores in an experiment on rote learning of paired associates (nonsense syllables).

Results. The trial, error, and total time scores (the second criterion trial excluded) are presented in Table 1. A comparison between the group means was made by one-way analyses of variance ($p>.10$ in the tests for homogeneity of variance). The outcome of this comparison is presented

Table 1.

Group	Trials		Errors		Total time	
	M	σ	*M*	σ	*M*	σ
A (NC-constant)	6.33	3.02	20.25	12.82	398.9	158.1
B (C-constant)	8.83	3.24	26.67	11.90	672.6	158.0
C (NC-variable)	13.25	3.88	50.08	17.29	842.0	206.2
D (C-variable)	9.17	2.64	26.75	8.14	631.9	129.7

Table 2.

Analysis of variance

Score	Source	SS	df	MS	F
Trials	Between groups	295.23	3	98.410	8.67
	Within groups	499.25	44	11.347	$p<.001$
	Total	794.48	47		
Errors	Between groups	6 198	3	2 066.0	11.29
	Within groups	8 053	44	183.0	$p<.001$
	Total	14 251	47		
Total time	Between groups	1 107 664	3	369 221	12.39
	Within groups	1 311 601	44	29 809	$p<.001$
	Total	2 419 265	47		

Differences between group means (df = 44)

		Trials			Errors		
		A	B	C	A	B	C
B	diff.	2.50			6.42		
	$t =$	*1.82*			*1.16*		
C	diff.	6.92	4.42		29.83	23.41	
	$t =$	*5.03***	*3.21***		*5.40***	*4.24***	
D	diff.	2.84	.34	4.08	6.50	.08	23.33
	$t =$	*2.07**	*<1*	*2.97***	*1.18*	*<1*	*4.22***

		Total time		
		A	B	C
B	diff.	273.7		
	$t =$	*3.88***		
C	diff.	443.1	169.4	
	$t =$	*6.29***	*2.40**	
D	diff.	233.0	40.7	210.1
	$t =$	*3.31***	*<1*	*2.98***

in Table 2. The differences between groups *B* and *D* (learning by the C method) are very slight and insignificant. Group *C* is significantly worse than any other group in all scores, while group *A* is significantly better than the other groups in time score. In trial score the difference between *A* and *D* is of borderline significance, while that between *A* and *B* does not quite reach significance ($p<.10$). In errors the differences between *A* on the one hand and *B* and *D* on the other are fairly small and not significant, although consistently in favour of *A*.

The results thus indicate, that in this experiment condition *A* was on the whole more economical than conditions *B* and *D* (between which the differences are negligible), while condition *C* was by far the most uneconomical one in all scores. The results are in full agreement with the prediction.

With respect to the total time score two points should be mentioned. The coefficients of variation are smaller than those for any other score. This is related to the fact that there was a fairly strong negative intra-group correlation between the trial and the average-time-per-trial scores: in group *A* $r = -.52$, in *B* $r = -.75$, in *C* $r = -.72$, and in *D* $r = -.71$; the faster a subject traced, the more trials tended to be required to learn the task. The 'average amount learned per unit time' was thus comparatively constant within each group. Secondly it may be pointed out that *inter*-group comparisons in terms of the total time score must be interpreted with caution, as each error increased the distance to be traced when the C method was used, while this was not the case when the NC method was used.

In order to check on the 'homogeneity' of the figures used in the experiment, the number of errors made in response to each figure was tabulated for groups *C* and *D*. With one exception the figures were found to be of almost equal difficulty.

A rough measure of the inter-trial variability was obtained in the following way. When the same wrong response was made to the same

figure on any two consecutive trials, the second error was termed repetitive. The percentage of repetitive errors (out of all errors) was computed for each S, excluding the first trial in each case (as no repetitive errors can be made on it). The average percentage of repetitive errors was 56.8 % in group *A* and only 42.6 % in *B*, the difference being significant ($p<.05$). This difference is consistent with previous findings, and will be discussed below.

Discussion. When the order of the figures was varied from trial to trial, and the associations between the figures and 'left' or 'right' consequently were formed relatively independently of each other, the correction method proved clearly superior. This result can comfortably be interpreted in terms of any present-day S-R model, and is hardly in need of further comment. With a constant order of presentation, however, the non-correction method was slightly but consistently more economical (some of the differences being significant). This is in agreement with the suggestion that the 'organization' of a constant choice sequence may be facilitated by use of the NC method. On the other hand, it is not in agreement with the prediction from the Guthrie-Voeks's theory discussed above; nor does a simple S-R 'chaining' model seem convenient for a discussion of the Ss' actual tracing behaviour under conditions *A* and *B*, presumably because psychologists developing this type of model have generally been more interested in the end-products of the learning process than in qualitative characteristics of the process itself. An analysis of the individual response sequences indicates that the most consistent difference between the 'ways' in which the subjects in groups *A* and *B* learned the task is of the following kind: Ss learning by the NC method tended to form a general impression or 'schema' of the main features of the whole sequence on the very first trials. This general impression gradually became more definite, and errors in it were often permanently corrected only when they had become localized in the total sequence. Consequently there was a comparatively great similarity between the patterns traced on consecutive trials, the percentage of repetitive errors being fairly high even when learning was fast. When the C method was used, the corrective movements tended to minimize the ('perceived') interdependence of the units, and rather than reacting to conceptions about the sequence 'as a whole', the Ss tended to learn the choices step by step, gradually connecting them together. The error distribution was predominantly progressive, the greatest number of errors being made on the last units. Such an interpretation of the individual records agrees well with the Ss' verbal comments during and after the experiments, and with the results from another comparable experiment (14). Although not

clear-cut in all cases, this difference between the characteristics of the performances in groups *A* and *B* is felt to have been sufficiently general to merit attention.

If the considerations presented in this paper are more generally valid, it seems likely that the NC method would be particularly favourable in cases where the serial organization of the whole sequence or of some aspects of it is of paramount importance for the performance. An example of such a situation is the learning of motor skills in which the part-activities are already in the S's behaviour repertoire, the sequential *timing* being the main feature to be learned – as frequently is the case, e. g., in 'industrial' skills and in sports.

SUMMARY

An attempt is made to put forward some distinctions relating to the correction and non-correction methods of learning, and some factors relevant to the relative economy of these methods are discussed. Particular attention is paid to human serial learning. It is predicted that in sensori-motor tasks of the paired associates type, the correction method should be superior when the order of presentation is varied from trial to trial, while the non-correction method may be more economical when a constant order of presentation is employed. An experiment is described in which this prediction is tested, and the results are found to confirm it. The experiment is discussed with reference to 'organization' in the learning of serial tasks.

REFERENCES

1. Buxton, C. E., & Bakan, M. B.: Correction vs. non-correction learning techniques as related to reminiscence in serial anticipation learning. *J. exp. Psychol.*, 1949, *39*, 338-341.
2. Hovland, C. I., & Weiss, W.: Transmission of information concerning concepts through positive and negative instances. *J. exp. Psychol.*, 1953, *45*, 175-182.
3. Hull, C. L., & Spence, K. W.: 'Correction' vs. 'non-correction' method of trial-and-error learning in rats. *J. comp. Psychol.*, 1938, *25*, 127-145.
4. Kosofski, S.: The effectiveness of positive and negative guidance as related to the degree of organization of the material to be learned. *J. gen. Psychol.*, 1952, *47*, 91-104.
5. McGeoch, J. A., & Irion, A. I.: *The psychology of human learning*. Rev. ed. New York: Longmans, Green, 1952.
6. Muenzinger, K. F., & Powloski, R. P.: Motivation in learning. X. Comparison of electric shock for correct turns in a corrective and a non-corrective situation. *J. exp. Psychol.*, 1951, *42*, 118-124.
7. Osgood, C. E.: *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
8. Seward, J. P.: An experimental analysis of maze discrimination. *J. comp. Psychol.* 1943, *35*, 17-27.
9. Thorndike, E. L.: *Human learning*. New York: Appleton-Century-Crofts, 1931.

10. Tolman, E. C.: The determiners of behaviour at a choice point. *Psychol. Rev.*, 1938, *45*, 1-41.
11. Ward, L. B.: Reminiscence and rote learning. *Psychol. Monogr., 1937, No. 220.*
12. Whitfield, J. W.: An experiment in problem solving. *Quart. J. exp. Psychol.*, 1951, *3*, 184-197.
13. Wischner, G. J.: The effect of punishment on discrimination learning in a non-correction situation. *J. exp. Psychol.*, 1947, *37*, 271-284.
14. von Wright, J. M.: An experimental study of human serial learning. D. Phil. Thesis, University of Oxford, 1955.